2023

Ballot Integrity Analysis: Unveiling Discrepancies and Patterns in Election Data

COMPARISON OF ORIGINAL ELECTION COUNT VS. RECOUNT USING BALLOT SIGNATURES

PHILLIP DAVIS

Contents

Introduction

In the recent election, questions have arisen regarding the integrity of the ballot counting process in certain precincts in Georgia. To shed light on these concerns, I embarked on a three-year detailed and thorough investigation, utilizing my skills as a mathematician and computer programmer to develop a custom software tool to aid in this process.

Every election represents the collective voice of the people, a democratic process that holds the potential to shape the future of a nation. The responsibility to ensure that this process is carried out with utmost integrity falls on the shoulders of those tasked with counting and verifying the ballots.

The Genesis of the Investigation

The 2020 presidential election marked a pivotal moment in America's history, with numerous events unfolding that questioned the integrity of the electoral process in Georgia. My interest was piqued when I stumbled upon a news segment that showcased a batch of around 100 duplicate ballots. Around the same time, Twitter, for reasons known only to them, decided to suspend my account for several years, disconnecting me from an essential source of information.

Rumors and whispers began circulating about possible discrepancies and anomalies in the Georgia ballot counting process, casting doubt on the sanctity of the democratic procedure that forms the bedrock of our nation.

These whispers turned into a clarion call for me, drawing me into an investigation fueled by a deep-seated passion for data analysis and a relentless pursuit of the truth. With my programming prowess as my ally and a sophisticated software tool that I had developed, I embarked on a journey. My mission was clear - to delve deep into the electoral abyss and compare the ballot images from the first count to those of the second count, in search of any irregularities that might have occurred.

Developing the Tool

The process of comparing thousands of ballot images is no small feat. It required a tool that could handle the complexity of the task and deliver precise results. Drawing on my years of experience as a computer programmer, I developed a software tool designed to load and display images side by side, facilitating easy comparison and analysis.

The software was built with the capability to import data and analyze it in a way that highlighted any discrepancies between the first and second counts. This was the key to unlocking the mysteries that lay hidden within the vast sea of ballot images.

Setting the Parameters

Before diving into the analysis, it was essential to set clear parameters that would guide the investigation. The first step was to establish a system of 'signatures' for each ballot. This involved identifying unique elements on each ballot, such as the precinct, ballot type, and the votes cast, and using them to create a signature that would serve as a fingerprint for each ballot.

The signatures were then used to group the ballots by precinct, creating distinct categories that could be analyzed separately. This was a crucial step, as it ensured that the comparison was carried out in a systematic and organized manner.

The Path Forward

With the tool in place and the parameters set, I was ready to embark on the path of discovery. Little did I know, the findings that awaited me would be both startling and revealing, shedding light on the intricacies of the ballot counting process and uncovering anomalies that had the potential to shake the very foundations of our democratic system.

Background in Ballot Analysis

Academic and Professional Foundation

I graduated from the University of Texas at Arlington in 1986, earning a Bachelor of Science in Mathematics with an emphasis on Computer Science. My career in the computer field spans over 38 years, reflecting a deep-seated expertise in technology and data analysis.

Specialization in Fingerprint and Image Analysis

My professional career into fingerprints and image analysis began about 30 years ago. I have been actively involved in the background check and fingerprint identification field, honing my skills in detailed image analysis and pattern recognition.

Political Involvement and Social Media Presence

My political engagement started in 2016, inspired by the election of President Trump. Despite being nearly 60, it was then that I registered to vote for the first time, drawn by the enthusiasm for a candidate I found refreshing.

- My foray into social media began with the creation of video memes, particularly following an incident where a meme creator was doxed by a news network.

- I initiated a Twitter account '@Mad_Liberals', which quickly amassed over 150,000 followers. My content mainly consisted of humorous mash-ups featuring President Trump and other political figures, with several videos reaching millions of views.

- Following the 2020 election, my account faced suspension for queries about signature verification. This incident not only fueled my curiosity about the election's integrity but also marked my deeper involvement in election analysis.

Motivated by a video showing duplicated ballots, I utilized my background in image analysis to explore ballot images from the Fulton election. This led to the creation of the "Ballot Finder" software, a tool designed to identify and analyze ballot duplications accurately.

Joining VoterGA, led by Garland Favorito, I gained access to a wider range of ballot images for analysis. This collaboration provided valuable resources and insights, enhancing the scope of my work.

Eventually, my collaboration with VoterGA concluded, leading to a new partnership with Joe Marolda. Together, we formed a Discord group with a few other like minded individuals to continue our ballot analysis efforts.

Our work faced unexpected challenges when Discord banned our account and deleted all our findings. Undeterred, Joe and I established a Slack channel, whimsically named 'Stacy Abrams Fan Club,' to mask our true nature and to continue our work in a more secure environment.

My start in ballot analysis is a testament to the interplay between technology, political interest, and social media dynamics. From a background in computer science and image analysis to navigating the complexities of social media and election integrity, my path reflects a blend of technical acumen and adaptive strategies in the ever-evolving landscape of political engagement and technological innovation.

Terminology

There are some terms that we use in the next section of the document.

HMPB – Hand Marked Paper Ballots

Hand Marked Paper Ballots (HMPB) are traditional voting ballots where voters manually fill in circles or boxes next to their chosen candidates or answers to referendum questions. These ballots are physically marked by the voter using a pen or pencil, typically by filling in a circle or completing an arrow next to the desired selection. HMPBs are particularly valuable in identifying duplicate ballots, as each voter's markings are unique. Even slight variations or marks outside the designated areas can serve as visual aids in the process of verifying the authenticity of each ballot and detecting duplicates. This individuality in marking style contributes to the integrity and security of the voting process.

BMD – Ballot Marking Device

A Ballot Marking Device (BMD) ballot is a type of modern voting ballot generated through a computerized system. Voters use a touchscreen interface to make their selections, and upon completion, the BMD prints out a paper ballot. This printed ballot includes the voter's choices in a format that is both human-readable and encoded in a QR code. BMD ballots are characterized by their uniform appearance; ballots with identical voting choices appear exactly the same, as they are machine-printed. This uniformity extends to the QR codes on these ballots, which precisely replicate for like-voted options, making them identical to other ballots with the same selections. However, this standardization presents challenges in election auditing, particularly in distinguishing between legitimately cast ballots and those that may have been scanned multiple times. The lack of unique, identifying marks, as found on hand-marked ballots, makes the detection of double-scanned BMD ballots more difficult. Despite these challenges, BMDs offer a digital method for voters to record their choices, combining electronic voting with a physical paper trail for tabulation and verification purposes.

DVS Number – Dominion Voting Machine Number

The DVS Number, short for Dominion Voting System Number, is a specific identifier associated with ballots processed by Dominion Voting Machines. This number plays a crucial role, particularly in the context of Ballot Marking Device (BMD) ballots. In BMD systems, the ballots are typically 'scrambled' or shuffled to obscure the order in which voters cast their ballots, aiming to maintain voter anonymity and ballot secrecy.

However, the DVS Number can be utilized to reverse this scrambling process, effectively restoring the original order of the ballots. This reordering is significant as it enables analysts to detect and study patterns in the voting data that would otherwise be concealed due to the scrambling process.

The methodology to reverse the ballot order using the DVS Number was developed by J. Alex Halderman, a notable figure in the field of election security and computer science. Further information and details about this process are available at https://dvsorder.org, providing a resource for those interested in understanding and applying this technique in ballot analysis.

MC1 – Machine Count 1

Machine Count 1 (MC1) refers to the initial and original counting of the ballots conducted for the 2020 presidential election. This count represents the first systematic tabulation of votes as collected from various voting methods, including Hand Marked Paper Ballots (HMPB) and Ballot Marking Device (BMD) ballots. MC1 is crucial as it establishes the baseline figures for the total votes each candidate received in the first tally of the election. The accuracy and integrity of MC1 are essential for a reliable and trustworthy electoral process, as it forms the primary record of voter preferences as initially registered and tabulated.

MC2 – Machine Count 2

Machine Count 2 (MC2) refers to the second recount of ballots that took place a few months following the initial count (MC1) in the 2020 presidential election. This recount serves as a critical process to verify and validate the results obtained from the original count. The purpose of MC2 is to ensure accuracy and reliability in the election outcome by providing a secondary, independent tally of the votes. This recount is especially significant in close races or in situations where discrepancies or procedural concerns were raised about the first count.

CVR – Cast Vote Record

The Cast Vote Record (CVR) is a database record generated within the Dominion Voting System, used to document and store detailed information about each individual vote cast in an election. The CVR includes essential data such as the tabulator number, batch identifier, precinct information, and polling ID. This comprehensive record serves as a digital footprint of each vote, capturing the specific details and context of the voting process. For analysis and review purposes, the CVR can be exported in various file formats, including JSON (JavaScript Object Notation) and CSV (Comma-Separated Values). Each vote in the election has a corresponding CVR.

JSON CVR – JSON formatted Cast Vote Record

The JSON CVR, or JSON formatted Cast Vote Record, refers to a digital representation of the Cast Vote Record (CVR) that is exported in the JSON (JavaScript Object Notation) format. JSON is a widely-used, text-based format for storing and transporting data, known for its simplicity and ease of use, particularly in programming and web development contexts.

In the realm of election analysis, the JSON CVR provides a structured and programmer-friendly way to access and manipulate the detailed voting data captured in the CVR. This format is especially beneficial for those involved in developing software tools and applications for analyzing and processing election data, as it allows for straightforward importing, parsing, and handling of complex voting records. The use of JSON format ensures that the data is not only easily accessible but also compatible with a wide range of programming environments and data processing tools.

CSV CVR – Comma Separated Value Cast Vote Record

The CSV CVR, or Comma Separated Value Cast Vote Record, refers to a format of the Cast Vote Record (CVR) that is exported as a CSV file. CSV is a simple file format used to store tabular data, such as

numbers and text, in plain text. Each line of the file is a data record, and each record consists of one or more fields, separated by commas.

Stray Ballots

Stray ballots" refer to ballots that are present in one set of voting records or counts but are absent in another, under circumstances where they are expected to be consistently accounted for across all records. They may arise due to various reasons such as administrative errors, handling mistakes, or issues in ballot processing.

Double-Counted ballots

This refers to ballots that have been scanned and recorded more than once in the vote tallying process, typically twice or sometimes even thrice. These repetitions occur when the same ballot is erroneously passed through the scanning machine multiple times. As a result, the same votes get counted more than once, which can lead to inaccuracies in the final vote tallies. In proper electoral procedures, these double scanned ballots should be identified and removed to ensure they do not unjustly influence the final vote count.

Duplicate ballots

This refers to a procedure used to address ballots that are not readable by the tabulating machines. In this process, a new ballot is created to exactly replicate the voter's original selections from the unreadable ballot. This duplication is conducted under strict guidelines and proper procedures to ensure the integrity and accuracy of the voter's intent. Duplicate ballots are a recognized and valid method for handling ballots that, due to damage, misprints, or other issues, cannot be processed by standard tabulation equipment. The creation of duplicate ballots is an essential practice in election administration and should not be misconstrued as a problematic or irregular action.

Introduction to Ballot Finder Software

Approximately three years ago, my journey in ballot analysis began, sparked by a video showcasing 100 duplicated ballots in Fulton County. This revelation, coupled with my extensive experience of nearly 30 years in fingerprint identification, positioned me uniquely to delve into the intricacies of ballot images. My expertise in scrutinizing image details drove me to explore this domain further.

Development of Ballot Finder

The inception of "Ballot Finder" marked the beginning of an extensive development process. This software, crafted in Microsoft .net C#, emerged from my desire to investigate ballot images.

Initial Features:

Image Comparison:

- The foundational functionality of the software allowed users to open and closely examine two images simultaneously, facilitating a side-by-side comparison. This feature included the ability to scroll and zoom into specific details of each image.

Navigation Tools:

- To enhance user experience, the software was equipped with next/previous image features, enabling seamless navigation through ballot images.

Lock Feature:

- A pivotal addition was the lock feature, which synchronized the movement of both images. This was instrumental in identifying and analyzing duplicated ballots that appeared in sequential order.

Reverse Feature:

- Frequently, double-scanned ballots were inserted upside down, resulting in a reversed sequence of the ballots. Activating the reverse sequence order button in the software enabled the lock feature to correctly sequence the second set of ballots in this reversed order.

QR Code Reading:

- An important enhancement to the "Ballot Finder" software was the integration of QR code reading functionality. This feature was specifically designed to handle Ballot Marking Device (BMD) ballots, which often contain QR codes encoding crucial voting information.

Advanced Development in "Ballot Finder" Software

As "Ballot Finder" evolved, its focus shifted towards harnessing more sophisticated image parsing capabilities. This progression in development led to significant enhancements in the software, making it an increasingly powerful tool in ballot analysis.

Key Developments in "Ballot Finder”

Ballot Signature Analysis

- The software was further refined to include the analysis of ballot signatures. This feature allows for the unique identification of a voter's intentions on a ballot, enabling the detection of duplicated ballots based on distinct voting patterns.

Importing and Converting Cast Vote Records

- "Ballot Finder" now supports the import of Cast Vote Records in both CSV and JSON formats. The software can convert this data into a signature format, facilitating a comprehensive analysis across different data types.

Jaccard Similarity Algorithm:

- A significant addition to the software is the implementation of the Jaccard Similarity algorithm. This algorithm is utilized to identify duplicated data within batches of ballots, aiding in the detection of double-scanned ballots.

Scan Time Analysis Feature

- The software now includes a feature for analyzing the scan times of ballots. This is particularly useful in identifying instances where double-scanned ballots occur in sequence, often because of scanner jams. By examining scan time discrepancies, the software can pinpoint cases of potential double capturing.

Exporting Data to Excel

- For enhanced usability and data presentation, "Ballot Finder" allows for the export of all analyzed data to Excel in a structured and formatted manner. This feature makes it easier to review, share, and further analyze the data.

Current State and Future Directions

"Ballot Finder" has become a testament to the power of continuous development and adaptation in the realm of ballot analysis. With each new feature, the software grows more capable of handling complex and varied tasks, demonstrating its efficacy in scrutinizing ballots with greater efficiency and accuracy. Moving forward, "Ballot Finder" will continue to evolve, driven by a commitment to uncovering and understanding the nuances of ballot integrity, ensuring that each phase of its development adds substantial value to the field of electoral analysis.

The Ballot Signatures

Introduction

This chapter outlines the methodology employed in the comprehensive analysis of ballots from the 2020 U.S. Presidential election in Georgia. The primary objective of this analysis was to conduct a comparative study between the results of the initial machine count and the subsequent machine recount, utilizing identical equipment as in the first count. An effective comparison was crucial, considering the observed discrepancies between the two counts. These discrepancies included instances of double-counted ballots, ballots unaccounted for in the original count, discrepancies in recount data, and ballots exclusively appearing in the second count.

A pivotal challenge in this analysis was devising a method to accurately juxtapose the two counts in the absence of unique identifiers like serial numbers or barcodes on the ballots. Furthermore, the investigation had to be adaptable to a variety of data formats obtained from different counties. This included ballot images, Comma-Separated Values (CSV) exports, and JavaScript Object Notation (JSON) exports of the voting databases. The latter two formats, collectively referred to as the Cast Vote Record (CVR), presented a unique set of challenges and opportunities for data analysis.

The focus of this chapter is to detail the strategies and processes utilized to transform these varied data forms into a coherent and analyzable format. This transformation was imperative to ensure a thorough and accurate comparison of the ballot counts, thereby underpinning the integrity and reliability of our analysis.

Ballot Signatures

Introduction

In the realm of election data analysis, the precise interpretation and translation of voter intent from ballot images into a comprehensible and analyzable format is crucial. This chapter, titled "Ballot Signatures," delves into the process of converting Cast Vote Records (CVRs) from ballot images into a structured and analyzable format. The focus here is on creating a unique representation of each voter's choices, referred to as a "ballot signature," which is instrumental in understanding voting patterns and behaviors.

Data Source and Initial Processing

Our primary data source comprises ballot images from diverse systems. Each system's output necessitates a standardized approach to ensure consistency in data interpretation. The CVR, located on the third page of the multi-page TIF ballot images, is pivotal in this process. It acts as a bridge between the physical, hand-marked paper ballots and the digital interpretation by the tabulator, which is programmed to recognize each race, candidate, and corresponding bubble placement.

Conversion to Excel Format

The initial task involves converting these CVRs into an Excel spreadsheet. This step is essential for transforming the data into an accessible and manipulable format. Each row in the spreadsheet corresponds to a single ballot, encompassing vital information such as the tabulator number, batch, ballot ID, and the date of scanning.

Creation of Ballot Vote Header

In our ongoing endeavor to refine the process of analyzing ballot images, the introduction of a 'Ballot Vote Header' represents a significant enhancement. This header serves as a preliminary identifier, providing crucial information about the type of ballot and its associated precinct details. The addition of this header to our ballot signatures systematizes and streamlines the analysis, making the signatures even more distinctive and informative.

Composition of the Signature Header

The Ballot Vote Header is composed of several key elements:

Ballot Type Indicator

- An initial letter signals the type of ballot: 'H' for Hand Marked Paper Ballot (HMPB) and 'B' for Ballot Marking Device (BMD) ballots.

Combo Code

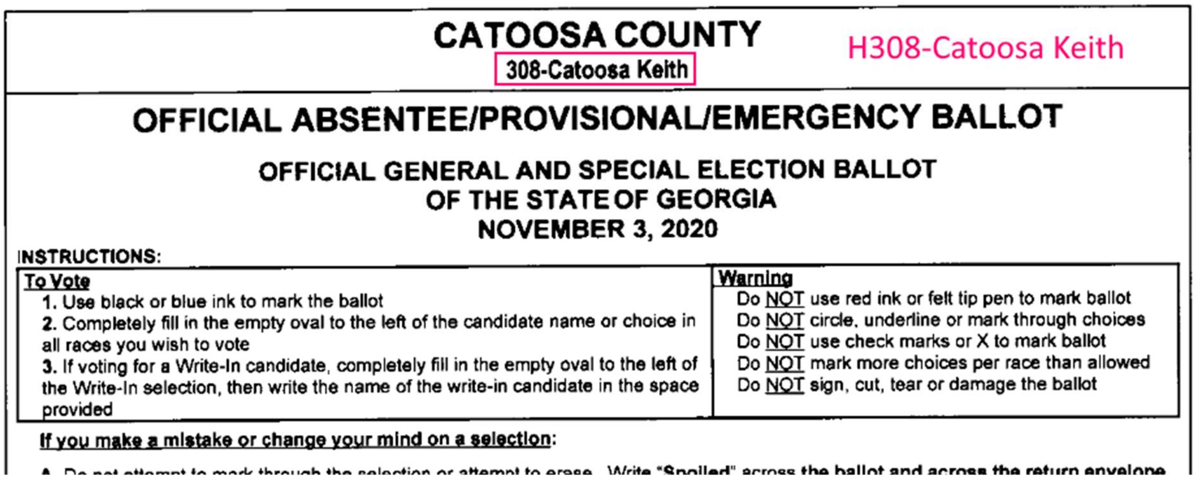
- Following the ballot type indicator, the Combo Code is added. This code, labeled as 'Ballot ID' on the third page of the ballot TIF images, further specifies the ballot's characteristics.

Precinct Identification

- The next component involves the precinct information. For single-precinct ballots, a 'Poll ID' is included in the Cast Vote Record on the third page.
- However, this Poll ID does not directly correlate with the precinct printed on the ballot. Therefore, an Optical Character Recognition (OCR) process is employed to read the combo/precinct information from the top of the first page and match it with the Poll ID. This step, known as 'Precinct Creation', is crucial but time-consuming due to the need for accuracy and correction of potential OCR errors.
- Once a complete precinct file is established, the Poll ID on the ballots is converted into a human-readable precinct code.

Example of a Complete Ballot Vote Header

For the given image, the Ballot Vote Header would appear as "H308-Catoosa Keith." This header indicates a Hand Marked Paper Ballot ('H'), with a combo code of '308', and the precinct named 'Catoosa Keith'.



CATOOSA COUNTY

308-Catoosa Keith

H308-Catoosa Keith

OFFICIAL ABSENTEE/PROVISIONAL/EMERGENCY BALLOT

OFFICIAL GENERAL AND SPECIAL ELECTION BALLOT
OF THE STATE OF GEORGIA
NOVEMBER 3, 2020

INSTRUCTIONS:

To Vote	Warning
1. Use black or blue ink to mark the ballot 2. Completely fill in the empty oval to the left of the candidate name or choice in all races you wish to vote 3. If voting for a Write-In candidate, completely fill in the empty oval to the left of the Write-In selection, then write the name of the write-in candidate in the space provided	Do NOT use red ink or felt tip pen to mark ballot Do NOT circle, underline or mark through choices Do NOT use check marks or X to mark ballot Do NOT mark more choices per race than allowed Do NOT sign, cut, tear or damage the ballot

If you make a mistake or change your mind on a selection:

Significance in the Analysis Process

The inclusion of the Ballot Vote Header transforms the ballot signature into a more precise and localized identifier. This aids in:

- Enhanced Precision: The header provides immediate context about the ballot type and its precinct, offering a more granular level of detail in the analysis.
- Efficient Sorting and Matching: With the header, the matching process can be more efficiently organized by precinct, facilitating a more streamlined analysis.
- Error Reduction: The process of creating the header, especially through the Precinct Creation step, minimizes inaccuracies and ensures a higher fidelity in the data.

The Ballot Vote Header is a pivotal addition to our ballot signature system. It not only augments the uniqueness of each signature but also enriches the data with essential contextual information. This advancement underscores our commitment to precision and thoroughness in the analysis of election data, ensuring that every aspect of the voter's choice is accurately captured and represented.

Composition of Signature Body

The cornerstone of our analysis lies in the creation of "ballot signatures." This innovative method involves:

Signature Composition:

- For each race, the first three letters of the candidate's name are concatenated to form a part of the signature. These segments are separated by colons for clarity and order.

Handling Special Cases

- **Blank Contests**: Races without a selection are marked as 'BLANK CONTEST' in the CVR. In such cases, "BLA" is used in the signature.
- **Overvotes**: Instances where more candidates are selected than allowed are marked as 'OVERVOTE'. Here, "OVE" is followed by the first three letters of each overvoted candidate in the signature.
- **Write-in Votes**: 'WRITE-IN' votes are accompanied by the candidate's name in the CVR. These are represented by "WRI" and the first three letters of the write-in candidate's name in the signature.

Ballot Type Specifics:

- Ballot Marking Machine (BMD) ballots, typically devoid of overvotes, lack this specific indicator in their signatures.
- Hand Marked Paper Ballots (HMPB) do not include the write-in candidate's name, thus only displaying 'WRI'.

00310_00007_000002.tif scanned at: 08:37:25 on 11/02/20.

Scanned on: ICC Tabulator: 310 Batch: 7
Ballot ID: 14

President of the United States
Donald J. Trump (I) (Rep)
US Senate (Perdue)
BLANK CONTEST
US Senate (Loeffler) - Special
Doug Collins (Rep)
Public Service Commission District 1
Jason Shaw (I) (Rep)
Public Service Commission District 4
Lauren Bubba McDonald, Jr. (I) (Rep)
US House District 14
Marjorie Taylor Greene (Rep)
State Senate District 53
Jeff Mullis (I) (Rep) Don:BLA:Dou:Lau:Mar:Jef:Dew:Chr:Jef:Tra:Gar:Gar:Jam:Joh:Ste:UNDRog:YES:YES:YES
State House District 3
Dewayne Hill (I) (Rep)
District Attorney - Lookout Mountain
Chris A. Arnt (Rep)
Probate Judge
Jeffrey D. Hullender (I) (Rep)
Clerk of Superior Court
Tracy Hullender Brown (I) (Rep)
Sheriff
Gary R. Sisk (I) (Rep)
Tax Commissioner
Gary Wayne Autry (I) (Rep)
Coroner
James T. "Jimmy" Spurling (Rep)
Chief Magistrate
John D. Gass (I) (Rep)
County Commission Chair
Steven M. Henry (I) (Rep)
Soil and Water - Catoosa
UNDER-VOTE BY: 2
Roger Bowman, Jr. (I)
Constitutional Amendment #1
YES
Constitutional Amendment #2
YES
Statewide Referendum A
YES

Significance of Ballot Signatures

The ballot signature concept is revolutionary in its ability to encapsulate the voter's pattern in a concise yet descriptive manner. It not only provides a unique identifier for each ballot but also maintains an ordered record of the voter's choices across different races. This semi-unique signature base facilitates comparative analysis, allowing us to identify patterns and anomalies in voting behaviors.

The process of creating ballot signatures represents a blend of technical precision and innovative thinking in election data analysis. By converting complex ballot data into a structured and analyzable format, we pave the way for more insightful and accurate interpretations of voter intentions and election outcomes. This chapter lays the groundwork for understanding and implementing this crucial aspect of election data analysis, emphasizing its importance in the broader context of electoral integrity and democracy.

Optical Character Recognition in Ballot Signature Analysis

The necessity for flawlessness in ballot signature extraction is paramount. Traditional optical character recognition (OCR) systems, while useful for many applications, have demonstrated limitations in this context. The primary issues encountered with commercial OCR solutions in ballot analysis include:

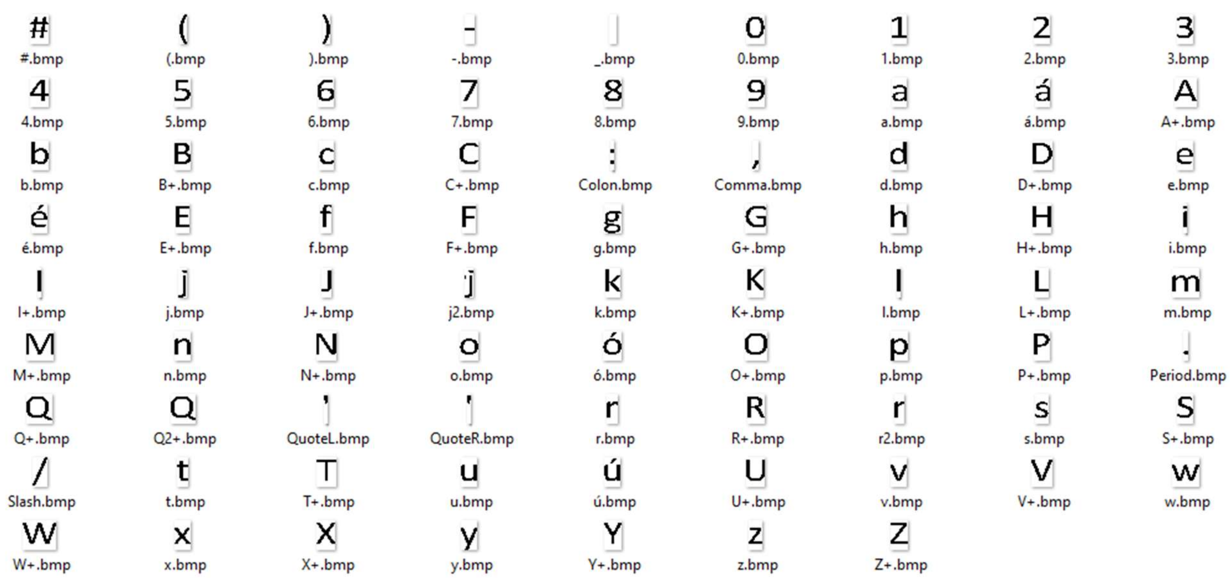
- Speed: Processing large volumes of ballot data, such as Fulton County's ballots, requires an immense amount of time, making the process inefficient.
- Accuracy: The reliability of OCR in accurately deciphering the data from the Cast Vote Record (CVR) on the TIF images is not sufficiently high for precise analysis.

Innovating with Pixel-Perfect OCR

To overcome these challenges, a more tailored approach was necessary. This led to the development of a unique OCR system specifically designed for the nature of the ballot data:

- Uniformity of the Cast Vote Record: Unlike the first two pages, the third page of the ballot, which contains the CVR, is generated by software, ensuring a consistent and uniform font. This uniformity is key, as it guarantees that each character, regardless of its position on the page, maintains the same pixel structure.
- Creation of a Character Dictionary: By collecting each character from the third page, a comprehensive dictionary of pixel-perfect font characters was developed. This dictionary allows for precise character recognition by matching each scanned character against the pre-defined pixel patterns.
- Efficiency and Accuracy: This customized approach significantly enhances both the speed and accuracy of data extraction. Characters are swiftly and accurately identified, vastly improving processing time and reliability.

- Handling Special Cases with Dipthongs: For instances where character separation is not clear-cut, 'dipthongs' – combinations of two or more characters – were created to address these unique scenarios.
- Font Library for Different Tabulators: Recognizing that different tabulators used varying font programs, a font library accommodating five different font bases was compiled. This adaptability is crucial for handling the variety of fonts encountered in the adjudication data.
- Success in Fulton County's Data: The effectiveness of this specialized OCR system was demonstrated in its ability to process the entirety of Fulton County's ballot database in just a few hours, a significant improvement over the days of processing required by standard OCR.



The development and implementation of this pixel-perfect OCR system marks a significant advancement in the field of ballot analysis. By creating a tailored solution that addresses the unique challenges of ballot data, we have significantly enhanced the accuracy and efficiency of the ballot signature extraction process. This innovation not only streamlines the analysis of large volumes of data but also ensures the integrity and reliability of the results, which are crucial for accurate electoral analysis and reporting.

The Comma Separated Cast Vote Record (CSV) in Ballot Signature Analysis

In the field of ballot analysis, data availability can significantly impact the scope and accuracy of the study. In cases like Fulton County, where not all ballot images are available, alternative methods are essential. The Comma Separated Cast Vote Record (CSV) provides a viable solution for such scenarios.

Utilizing CSV for Signature Generation

The CSV, derived from the voting software as a database export, offers a rich dataset that can be pivotal in generating ballot signatures. This method is particularly useful when full original and recount images are not available for all counties.

Key Features of the CSV Approach:

- Binary-Like Data Representation: Each vote is recorded as '1' for yes and '0' for no, resembling a binary format. This structure, while not human-readable in its raw form, forms the basis for generating readable ballot signatures.
- Adapting Ballot Image Analysis Techniques: The same techniques applied to the third page of the ballot images can be employed with the CSV data. This consistency ensures uniformity in the signature generation process.
- Identification of Ballot Types: The 'Counting Group' column in the CSV helps distinguish between Hand Marked Paper Ballots (HMPB) and Ballot Marking Device (BMD) ballots, with 'H' indicating Absentee By Mail ballots and 'B' for Election Day/Early Vote ballots.
- Precinct and Combo Information: The 'PrecinctPortion' field in the CSV already contains formatted combo/precinct information, streamlining the signature creation process.
- Capturing Voting Patterns: The signatures are built by identifying '1' values in each column, representing votes for candidates. The first three letters of each voted candidate's name are added to the signature, separated by colons for clarity.
- Handling Special Voting Scenarios: The 'Vote for' field is crucial for identifying overvotes (OVE), undervotes (UND), blank ballots (BLA), and write-in ballots (WRI).

Comparison of Signatures from Different Sources

The signatures generated from both the ballot images and the CSV tend to align well, showcasing the robustness of the signature creation methodology. However, it is important to note the distinction in data types:

- Adjudicated vs. Original Voting Intention: CSV records often include adjudicated data, whereas image ballots reflect the original voting intention. Caution is advised when comparing signatures from these different sources.
- Comparative Analysis: While it is preferable to compare like sources, with careful analysis, it is possible to draw meaningful comparisons between signatures derived from different sources.

The use of Comma Separated Cast Vote Records in ballot signature analysis provides a valuable alternative when ballot images are not fully available. This approach not only compensates for the lack of visual data but also maintains the integrity and consistency of the signature generation process. As demonstrated, this method can successfully be integrated into the broader framework of ballot analysis, ensuring comprehensive and accurate election data interpretation even in the face of data limitations.

	A	B	C	D	E	F	G	I	J	K	L	M	N	O	P	Q	
1	Bacon Nov	5.5.12.1															
2								President	President	President	President	President	President	President	President	President	P
3								Donald J. T	Joseph R.	Jo Jorgens	Write-in	Loren Coll	Gloria La F	Barbara Be	Brian Carr	David Byrn	D
4	CvrNumbe	Tabulator	BatchId	RecordId	Imprinted	CountingGroup	PrecinctPortion	BallotType									
5	1	70	1	1	70-1-1	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
6	2	70	1	2	70-1-2	Absentee by Mail	005-Douglas	1	0	0	0	0	0	0	0	0	
7	3	70	1	3	70-1-3	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
8	4	70	1	4	70-1-4	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
9	5	70	1	5	70-1-5	Absentee by Mail	007-Douglas	0	1	0	0	0	0	0	0	0	
10	6	70	1	6	70-1-6	Absentee by Mail	005-Douglas	1	0	0	0	0	0	0	0	0	
11	7	70	1	7	70-1-7	Absentee by Mail	005-Douglas	1	0	0	0	0	0	0	0	0	
12	8	70	1	8	70-1-8	Absentee by Mail	007-Douglas	0	1	0	0	0	0	0	0	0	
13	9	70	1	9	70-1-9	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
14	10	70	1	10	70-1-10	Absentee by Mail	005-Douglas	1	0	0	0	0	0	0	0	0	
15	11	70	1	11	70-1-11	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
16	12	70	1	12	70-1-12	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
17	13	70	1	13	70-1-13	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
18	14	70	1	14	70-1-14	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
19	15	70	1	15	70-1-15	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	
20	16	70	1	16	70-1-16	Absentee by Mail	007-Douglas	0	1	0	0	0	0	0	0	0	
21	17	70	1	17	70-1-17	Absentee by Mail	007-Douglas	0	0	0	0	0	0	0	0	0	
22	18	70	1	18	70-1-18	Absentee by Mail	007-Douglas	0	1	0	0	0	0	0	0	0	
23	19	70	1	19	70-1-19	Absentee by Mail	007-Douglas	0	1	0	0	0	0	0	0	0	
24	20	70	1	20	70-1-20	Absentee by Mail	004-Douglas	1	0	0	0	0	0	0	0	0	

The JSON Cast Vote Records

In the evolving landscape of ballot analysis, different data formats offer varied insights and challenges. This chapter focuses on the utilization of JSON (JavaScript Object Notation) format for analyzing cast vote records. JSON, a widely used format for storing and transmitting data, presents a unique opportunity for ballot signature analysis, particularly when other forms of data, like ballot images, are unavailable or incomplete.

JSON in Ballot Signature Analysis

Advantages of JSON Format:

- Direct Data Dump: The JSON files used in this analysis are direct exports from the cast vote records in the database, ensuring a high degree of accuracy and completeness in the data.
- Inclusion of Fill Amount: A distinctive feature of the JSON format in this context is the inclusion of the fill amount per bubble, offering an additional layer of data that can be critical in certain analyses.

Limitations and Workarounds:

- Lack of Write-in Candidate Names: Unlike the Ballot Image format, the JSON data does not include the names of write-in candidates. However, this is mitigated by the fact that ballot images also rarely include these names, allowing for consistent comparison using just the 'WRI' designation.
- Handling Special Cases with Outstack Conditions: The JSON data introduces the concept of Outstack Conditions, which are critical in identifying overvotes, undervotes, blank ballots, and write-in ballots. This feature enhances the ability to accurately categorize and analyze ballots.

Comparison with CSV Data

The JSON and CSV data formats, while different in structure, both originate from the same underlying database. This similarity ensures that the methods and logic applied in the conversion and analysis process remain consistent across both formats. The key difference lies in the data representation, with JSON providing a more hierarchical and structured format compared to the flat structure of CSV.

Application in Fulton Ballot Analysis

In the case of Fulton County, where most original ballot images were unavailable, the JSON Cast Vote Record format proved to be invaluable. By leveraging this data format, a comprehensive analysis of the Fulton ballots was possible, despite the absence of physical ballot images. This highlights the flexibility and adaptability of the analysis process in accommodating different data formats to achieve a complete and accurate understanding of voter intentions.

Conclusion

The exploration of JSON Cast Vote Records in ballot signature analysis underscores the importance of versatile data handling in election analysis. This format, with its unique features and compatibility with other data types, enables a thorough and nuanced examination of voting patterns. As demonstrated in the Fulton County case, the ability to adapt to different data formats is crucial in ensuring a comprehensive analysis, especially when dealing with incomplete or missing data sources.

The Aberration System

Introduction

In the pursuit of accurately identifying voter intent and comparing ballot counts across original tallies and recounts, a robust methodology is essential. This chapter introduces "The Aberration System," a sophisticated approach used in the 'Ballot Finder' software to analyze discrepancies in ballot signatures from the original count to the recount.

Methodology for Comparing Counts

The software initiates its process by prompting the user to load Excel signature files, which are created either from TIF ballot images or cast vote records. Additionally, it requests the original and recount signature files, along with an 'aberration input file'—a document encompassing a list of manual corrections that will be developed later on. Upon starting the process, the software's algorithm runs for a few minutes, diligently working to identify any discrepancies in the occurrence of ballot signatures.

The analysis within the software is conducted on a precinct-by-precinct basis, comparing the frequency of each signature between the original count and the recount. In order to streamline the process, signatures that exhibit matching frequencies across both counts are immediately excluded from further scrutiny. Attention is then directed to signatures that display variance in their occurrence counts, which may signal potential irregularities.

Importantly, the software distinguishes between Ballot Marking Device (BMD) ballots and Hand Marked Paper Ballots (HMPB), ensuring that comparisons are only made within the same type of ballots. This distinction is maintained by the first letter of the signature – 'B' for BMD and 'H' for HMPB – preventing any cross-type signature matching.

Upon completion of this analysis, an 'aberration' Excel spreadsheet is generated. This output file is comprehensive and includes several key fields for in-depth analysis:

- Aberration Type: Categorized as Stray, Match, or Doubled.
- Precinct/Combo: Identifying the specific precinct and combo code.
- Tabulator/Batch/Ballot: Details of the tabulator used, batch number, and ballot identifier.
- Occurrences: The frequency of each signature’s appearance.
- Gap: A metric indicating the interval between the current ballot and the preceding one in the sequence.
- Signatures: The specific signature associated with each ballot.

These fields collectively provide a detailed overview of each ballot's process through the counting process, highlighting areas that require further investigation or verification.

Discrepancy Types

These occurrence discrepancies can arise due to several reasons:

1. Adjudication Variations: In some cases, discrepancies result from the adjudication process, where ballots are reviewed and potentially modified to reflect the voter's intent more accurately. This can include correcting overvotes, disqualifying write-in votes, or reinstating omitted candidates.
2. Double or Triple Scanning: Another common source of discrepancy is the multiple scanning of ballots. Double-scanned ballots result in an additional occurrence, while triple scans lead to two extra occurrences.
3. Missing Ballots: Variations in ballot occurrences can also stem from stray or missing ballots, where batches of ballots are included in one count but not the other.

Count Error and Stray Error

A specific type of count error, known as a 'stray' error, occurs when a ballot appears in one count but not the other. Highlighting these errors through color-coding aids in the quick identification of large batches of stray ballots.

Outputting the Aberration File

The culmination of this process is the creation of the 'Aberration File,' an Excel spreadsheet that lists all count and stray errors in an organized manner - by tabulator, batch, and ballot order. This organization allows for the easy identification of patterns, such as large batches of duplicated or added ballots.

Each precinct's count and stray issues are then sorted into their individual tabs within the spreadsheet.

Armed with the Aberration File, the next step involves a manual review process to delve deeper into the identified discrepancies, aiming to uncover the underlying causes and implications of these anomalies.

Identifying Matching Ballots

The aberration file provides a comprehensive list of signatures with discrepancies in their occurrence counts. The key objective now is to determine the reasons behind these variances.

One of the most frequent and benign reasons for occurrence issues is adjudication, particularly in cases where voters use check marks or small 'x' marks instead of fully filling in the circles as instructed. Such voting methods often lead to adjudication and/or tabulator interpretation errors. In ballot image analysis, an adjudicated ballot typically appears as a 'stray' single signature in one count and a 'mismatch' error in the other.

The first step in addressing adjudication involves thoroughly reviewing each precinct’s tab in the adjudication Excel report. It is advisable to complete adjudication for one precinct before moving on to the entire document.

Procedure for Hand Marked Paper Ballots (HMPB)

When encountering a stray ballot in one count and a mismatch in the other, the corresponding file name of the stray ballot should be copied into Ballot Finder to view the image. Subsequently, all files with a similar signature on the other count should also be copied for comparison.

With Ballot Finder's next/previous buttons, users can systematically examine each potential match to the stray ballot, looking for unique marks or circles characteristic of hand-marked ballots. Pressing the next button or hitting the F3 button will quickly advance to the next button. To examine a hundred ballots against a candidate typically takes under a minute.

Upon finding a match, the 'match' button in Ballot Finder should be clicked to record the pairing. The user will then go into the Aberration Input file, and paste the results into the Excel spreadsheet. This adds a new 'Match' record with both matching ballots in the Aberration Input File.

Procedure for Ballot Marking Device (BMD) Ballots

For Ballot Marking Device (BMD) ballots, the process of comparison is distinctly different due to the lack of unique identifiers on each ballot. The primary focus is on ensuring that the QR codes on the respective ballots align precisely. However, this often presents a challenge as there may not always be a clear 'matching print', leading to situations where users have to rely on an educated guess to determine which ballot is the duplicate.

In instances where previously selected ballots have established a pattern of aberrations predominantly associated with a specific tabulator, users might choose a ballot from this tabulator as the likely duplicate. Alternatively, a common approach is to select the ballot with the lowest number in the sequence as the probable match.

Given the inherent limitations in matching BMD ballots, it's crucial to exercise caution and considerate judgment when making these selections. The absence of definitive matching criteria means that any decision to pair ballots as duplicates must be approached with an understanding of the potential for error, ensuring that the process remains as accurate and reliable as possible within these constraints.

Resolving Aberrations

The process of aberrating ballots involves a methodical, precinct-by-precinct approach to ensure accuracy and thoroughness in the analysis. Utilizing the Ballot Finder software is essential in this process, as it provides a visual means to compare and match hand-marked ballots effectively. All identified matches should be diligently documented in the Aberration Input File. It's important to be prepared for this to be a time-intensive process, particularly in precincts where there are numerous aberration strays to resolve.

The Stray Finder algorithm and methodology are structured to facilitate iterative analysis and refinement. As you identify and categorize ballots as 'match', 'duplicate', or 'stray', and then rerun the analysis, the system dynamically updates the occurrence counts of each signature. Each time a ballot is categorized, the occurrence count for that specific signature decreases by one. Consequently, as the process progresses and more ballots are categorized, the number of signatures with mismatched occurrences starts to diminish. Signatures with now-equal occurrences across counts cease to appear in

the report. This iterative nature of the process is crucial, with the goal of running it multiple times until all aberrations are effectively resolved.

For every signature that displays a mismatch in occurrences, the system prints out every potential instance of its use. For example, if a particular ballot signature appears 50 times instead of the expected 49, all 50 instances are displayed. This approach allows human analysts to discern patterns more easily. Once the additional, or 50th, signature is addressed—either through matching it to another ballot or identifying it as a stray—the system automatically resolves and removes the other 49 instances from the report. This method ensures a thorough and systematic approach to resolving discrepancies in ballot occurrences.

The Gap Count

In the process of displaying potential locations for mismatched ballots, the software includes a critical feature: the 'gap' count, which indicates the number of records between successive ballots. If a series of ballots is displayed in incremental order, the gap count will be 1 for each. To draw attention to sequences of ballots that are in sequential order, the software highlights these gap counts in red. This visual cue is especially useful for the analyst to quickly identify that the ballots are sequentially ordered.

The significance of this feature becomes more apparent in the case of entire batches of stray ballots or series of double-scanned ballots. In such scenarios, the gap count will consistently show as 1 and will be displayed in red, providing a clear and distinct signal to the user.

The visual distinction between a batch of stray ballots and double-scanned ballots is made evident through this system. For example, a sequence of missing ballots typically exhibits multiple single occurrences of stray ballots. This pattern causes the entire sequence to be highlighted in a yellow font, signifying stray ballots.

Stray		1	1	00580_00015_000001	B083-CO:Jos:Jon:Kan:Eli:Dan:Kev:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:YES:YES:YES
Stray	7	1	1	00580_00015_000008	B083-CO:Jos:Sha:All:Mar
Stray	51	1	1	00580_00015_000059	B078-CO:Jos:Sha:EdT:Rob:Nat:Kev:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:YES:NO_:NO_
Stray	1	1	1	00580_00015_000060	B078-CO:Jos:Jon:Ric:Jas:Dan:Mar:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:NO_:NO_:YES
Stray	1	1	1	00580_00015_000061	B078-CO:Jos:Jon:Rap:Jas:Dan:Mar:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:NO_:NO_:YES
Stray	1	1	1	00580_00015_000062	B078-CO:Jos:Jon:Kan:Rob:Lau:Mar:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:NO_:NO_:YES
Stray	1	1	1	00580_00015_000063	B078-CO:Jos:Sha:Tam:Rob:Dan:Mar:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:YES:YES:YES
Stray	1	1	1	00580_00015_000064	B083-CO:Jos:Dav:Rap:Jas:Nat:Mar:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:YES:NO_:NO_
Stray	1	1	1	00580_00015_000065	B083-CO:Jos:Sha:Rap:Rob:Lau:BLA:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:YES:YES:NO_
Stray	1	1	1	00580_00015_000066	B078-CO:Jos:Sha:AlB:Rob:Lau:Mar:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:NO_:NO_:YES:YES
Stray	1	1	1	00580_00015_000067	B078-CO:Jos:Jon:AWa:Rob:Dan:Kev:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:YES:YES:YES:YES
Stray	3	1	1	00580_00015_000070	B078-CO:JoJ:Dav:Kan:Eli:Nat:Mar:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:NO_:YES:YES:NO_
Stray	1	1	1	00580_00015_000071	B083-CO:JoJ:Dav:EdT:Eli:Lau:Kev:Chu:Jas:Ber:She:BBa:Sco:Dan:Gre:Jev:Ron:Car:BLA:NO_:YES:YES:NO_

Conversely, double-scanned ballots, by their very nature, are never categorized as stray since they occur more than once. Therefore, these sequences are displayed in blue, aligning with the 'match' category. This color-coded system aids in distinguishing double-scanned ballots from stray ballots.

Count	1	2	1	00116_00028_000069	B201-Yatesville:Jos:Jon:Rap:Jas:Dan:Val:BLA:Chr:BLA:Dan:Dav:Jam:Nor:Lor:Sid:YES:YES:NO_
Count	1	17	1	00116_00028_000070	B201-Town:Jos:Jon:Rap:Rob:Dan:Val:Joh:Chr:Ter:Dan:Dav:Jer:Deh:Lor:Sid:YES:YES:YES
Count	1	2	1	00116_00028_000071	B203-Redbone:Jos:Jon:Deb:Rob:BLA:BLA:BLA:Chr:BLA:BLA:BLA:BLA:Nor:Jam:Sid:YES:YES:YES
Count	1	3	1	00116_00028_000072	B207-Redbone:Jos:Jon:Rap:Rob:Dan:Val:BLA:Chr:BLA:BLA:BLA:Jer:Deh:BLA:YES:NO_:YES
Count	1	2	1	00116_00028_000073	B209-Redbone:Jos:Jon:Rap:Rob:Dan:Val:BLA:Chr:BLA:BLA:BLA:Jer:Deh:Sid:YES:YES:NO_
Count	1	7	1	00116_00028_000074	B207-Redbone:Jos:Jon:Rap:Rob:Dan:Val:BLA:Chr:BLA:BLA:BLA:Jer:Deh:BLA:YES:YES:YES
Count	1	4	1	00116_00028_000075	B207-Redbone:Jos:Jon:Rap:Rob:Dan:Val:BLA:Chr:BLA:BLA:BLA:Jer:Deh:BLA:NO_:NO_:YES
Count	1	3	1	00116_00028_000076	B232-Town:Jos:Jon:Deb:Rob:Dan:Val:BLA:Chr:BLA:BLA:BLA:Jer:Deh:BLA:BLA:BLA:BLA
Count	1	10	1	00116_00028_000077	B213-Town:Jos:Jon:Deb:Rob:Dan:Val:Joh:Chr:Ter:Dan:Dav:Jer:Deh:Lor:Sid:YES:YES:YES
Count	1	3	1	00116_00028_000078	B201-Town:Jos:Jon:Rap:Rob:Dan:Val:Joh:Bet:Ter:Dan:Dav:Jer:Deh:Lor:Sid:YES:NO_:NO_
Count	1	5	1	00116_00028_000079	B215-Town:Jos:Jon:Rap:Rob:Dan:Val:Joh:Chr:Ter:Dan:Dav:Jer:Deh:Lor:Sid:NO_:YES:YES
Count	1	2	1	00116_00028_000080	B237-Town:Jos:Jon:Deb:BLA:BLA:BLA:BLA:BLA:BLA:BLA:BLA:BLA:BLA:BLA:BLA:BLA:BLA
Count	1	2	1	00116_00028_000081	B223-Town:Jos:Jon:EdT:Rob:Dan:Val:BLA:Chr:BLA:BLA:BLA:Jer:Deh:BLA:BLA:NO_:YES:YES
Count	1	2	1	00116_00028_000082	B220-Town:Jos:Jon:Rap:Rob:Dan:Val:BLA:Chr:BLA:Dan:BLA:Jer:Deh:BLA:BLA:YES:YES:NO_
Count	1	2	1	00116_00028_000083	B223-Town:Jos:Jon:Rap:Rob:Dan:Val:BLA:Chr:BLA:Dan:BLA:Jer:Deh:BLA:BLA:YES:YES:NO_
Count	1	6	1	00116_00028_000084	B204-Redbone:Jos:Jon:Rap:Rob:Dan:Val:Joh:Chr:Ter:Dan:Dav:Jer:Deh:Sid:NO_:YES:YES
Count	1	2	1	00116_00028_000085	B211-Town:Jos:Jon:Deb:Rob:Dan:Val:Joh:Chr:Ter:Dan:Dav:Jer:Deh:Sid:YES:NO_:YES
Count	1	2	1	00116_00028_000086	B201-Town:Jos:Jon:Rap:Rob:Dan:Val:BLA:Chr:BLA:BLA:BLA:Jer:Deh:Lor:BLA:YES:YES:BLA
Count	1	2	1	00116_00028_000087	B202-Yatesville:Jos:Jon:Mat:Rob:Dan:Val:Joh:Chr:Ter:Dan:Dav:Jer:Deh:Lor:Sid:YES:YES:YES

Double Scanned Ballots

The analysis of double scanned ballots in election data involves a multi-step process using the Ballot Finder software, which includes several features designed to facilitate the identification and confirmation of these discrepancies.

- Identification of Double Scanned Batches: Initially, both the original set of double scanned batches and the corresponding original set of ballots are flagged. This identification is crucial in understanding the extent of duplication within a specific batch.

Using Ballot Finder for Confirmation:

- The list of potentially double scanned ballot names is entered into the 'batch file open' feature of Ballot Finder. The analyst places the full list of suspected duplicate ballots in one window and the target list of ballots in another.
- Utilizing the 'lock' feature of the software, both sets of potentially double scanned ballots can be viewed simultaneously, allowing for visual confirmation of duplicates.

Reverse Order Matching: Often, double scanned ballots are placed in the scanner upside down, complicating the matching process. The 'reverse order' button in Ballot Finder addresses this challenge by moving one set of ballots forward and the corresponding set backward, facilitating easier comparison.

Randomized Duplication Patterns: In many cases, the order of the duplicated batch is randomized, deviating from its original sequence. Sometimes, batches are duplicated by scanning small groups of previous ballots in both normal and reversed order. This randomness necessitates careful analysis to identify true matches.

Speculation on Duplication Intent: The randomness in the duplication process raises questions about whether such occurrences are accidental or intentional attempts to conceal duplication.

Additional Duplication Checks

The 'Jaccard' algorithm is used to check if an entire batch is a duplicate of another, even if the order is randomized.

Another feature performs a detailed check through runs of signatures, seeking similar patterns in other batches. While more time-consuming, this method can be highly effective, especially when only one set of cast vote records is available. The results is a list of doubled prints that is ready to copy into your aberration report.

Identifying 'Jam Doubles'

'Jam doubles' occur when a scanner jams, and the operator rescans the last few ballots. These often appear as small groups of two to four duplicated ballots in sequence, or occasionally as a single ballot immediately following its first scan.

The stray finder feature is particularly adept at identifying jam doubles, indicating an additional occurrence of a ballot without a corresponding match. Careful examination of the sequence helps determine if other ballots in the batch were also duplicated.

Through these methods and tools, the Ballot Finder software provides a comprehensive approach to identifying and analyzing double scanned ballots, offering insights into their nature and the processes leading to their occurrence.

Completing The Aberration Process

As you diligently work through the process of identifying matches, strays, and duplicates using "The Aberration System," you will reach a point where the initial 'ToDo' tab in your Excel spreadsheet begins to deplete. This tab, initially filled with potential discrepancies to be resolved, will gradually clear up as you validate each visual match you enter. The completion of this tab signifies the end of the primary matching process.

Resolution of Precinct Tabs

Each precinct tab in the spreadsheet represents a unique set of aberrations to be addressed. As you resolve the issues in each precinct, these tabs will also disappear. Eventually, you will be left with two main tabs: 'MC1 Aberrations' and 'MC2 Aberrations'.

The Final Results

The two tabs named MC1 and MC2 contain the culmination of your analysis:

- Stray Ballots: A comprehensive list of all the stray ballots you have identified throughout the process.
- Duplicate Ballots: A detailed listing of all duplicate ballots, including the file names of the original ballots for cross-reference. The software enhances the verification process by displaying both sets

of signatures. When a genuine match is confirmed, the corresponding signature turns green, providing an additional layer of assurance that the identified doubles are accurate.

With the MC1 and MC2 aberration tabs finalized, your stray analysis is complete. The duration of this process can vary significantly, ranging from a few hours for less complex counties to several months for more intricate cases. For instance, resolving Fulton County's aberrations took over five months, including multiple restarts due to its complexity. This variance in time reflects the diverse challenges and nuances encountered in different counties, underscoring the thorough nature of the aberration analysis process.

The Counties of Georgia

Introduction

In our comprehensive review of the 2020 election data from Georgia, our analysis was focused on those counties where we had access to both the first machine count (MC1) and the second machine count (MC2) Cast Vote Records. This availability of data was a crucial criterion for our stray analysis, allowing us to compare and contrast the voting records effectively. However, this requirement limited our scope to less than 50 of the 159 counties in Georgia, representing just a fraction of the state's total.

From this subset of counties, more than a third exhibited some form of counting anomaly, which we will detail in the individual sections for each county. These discrepancies range in nature and magnitude, shedding light on various aspects of the counting process. Before delving into these specific cases, it's important to acknowledge the counties where our analysis found no issues. The following list outlines those counties that, according to our algorithm and criteria, displayed a smooth and consistent counting process without any significant anomalies. This serves as a baseline, offering a contrast to the more complex situations we encountered in other counties.

Counties with No Aberrations

- Bacon
- Bartow
- Baker
- Barrow
- Bleckley
- Candler
- Chattooga
- Colquitt
- Columbia
- Cook
- Dade
- Gordon
- Irwin
- Lowndes
- Lumpkin
- Madison
- Marion
- Mitchell
- Morgan
- Oconne
- Pierce
- Pike
- Rockdale
- Schley
- Stewart
- Talbot
- Tattnall
- Warren
- Wilcox

Counties With Aberrations

In our detailed examination of the 2020 election data, we encountered various counties exhibiting notable discrepancies in their ballot counts. The following sections present an overview of counties that displayed significant instances of double ballots or stray ballot aberrations. These aberrations, ranging from minor irregularities to more pronounced anomalies, provide insight into the complexities and challenges inherent in the ballot counting and verification processes.

County	Double MC1	Stray MC1	Double MC2	Stray MC2	Total
Bibb	0	103	0	22	125
Bullock	0	0	5	1	6
Catoosa	0	0	0	28	28
Chatham	185	32	1407	27	1651
Cherokee	125	0	22	127	274
Cobb*	0	482	399	28	909
DeKalb	0	129	0	160	289
Floyd	4	12	0	2923	2939
Forsyth	0	0	23	0	23
Fulton**	391	3284	3920	1413	9008
Grady	50	0	0	0	50
Gwinnett	449	963	151	575	2138
Houston	0	4	0	26	30
Muscogee	50	434	455	440	1379
Pickens	0	1	0	179	180
Upson	0	101	100	0	201
Walton	0	0	0	284	284
Ware	100	2	0	51	153
Whitfield	1	8	0	30	39
Total	1355	5554	6482	6135	19.706

* Indicates that we do not have the full set of TIF images to complete the aberration study

** indicates that the county is still not fully resolved yet.

Bacon County

Data Sources and Ballot Counts

In Bacon County, our analysis is grounded in data obtained from multiple sources:

- MC1 TIF ballot images
- MC1 CSV Cast Vote Record: Bacon Nov 2020 General (4,680 records)
- MC2 CSV Cast Vote Record: Bacon Nov 2020 Recount (4,680 records)

Initial Observations and Misinterpretations

Bacon County serves as a prime example of the necessity for thorough analysis. At first glance, certain aspects of the data may raise suspicions. For instance, a comparative analysis between the two datasets revealed 23 ballots that seemed unusual. These ballots, all from the BMD (Ballot Marking Device), only had the presidential candidate 'Trump' selected, leaving other races blank. These 23 ballots appeared sequentially at the end of the count, prompting initial concerns about their legitimacy.

Uncovering the Truth

However, upon deeper investigation, it was found that these 23 ballots were legitimate and resulted from a procedural necessity. It appears that the original ballots could not be properly scanned. During the recount, which focused solely on the presidential race, the down-ballot races were not a consideration. The election personnel, adhering to protocol, correctly remade these 23 unscannable ballots, duplicating them and recasting the votes.

Areas for Improvement

Despite Bacon County's overall adherence to proper procedures, there were areas that required further scrutiny:

- Mismatch in Precincts: The precincts for the duplicate ballots did not align with those of the original ballots. While the original ballots spanned five different precincts, the duplicates were limited to three. This discrepancy necessitated additional analysis to clarify the situation.
- Adjudication of Write-in Votes: In the initial count, many write-in votes were changed to 'blank', indicating that the BMD ballots underwent adjudication. This process likely involved the removal of candidates not meeting qualifications, an additional manual step essential for accurate ballot analysis.

Conclusion

Bacon County's case highlights the importance of comprehensive ballot analysis. While the initial data review suggested anomalies, a detailed examination revealed that the county efficiently managed challenges such as unscannable ballots. The identified areas for improvement emphasize the need for attention to detail in every aspect of ballot handling and analysis, ensuring the integrity of the electoral process.

Bartow County

For Bartow County, the source of the information relies on images and data files from the following sources.

- MC1 TIF Ballot Images
- MC1 JSON Cast Vote Record: Bartow Nov 2020 General
- MC1 CSV Cast Vote Record (50,678 records)
- MC2 JSON Cast Vote Record: Bartow Nov 2020 Recount
- MC2 CSV Cast Vote Record (50,675 records)

Context and Overview

In the analysis of Bartow County's ballots, we encountered a scenario akin to that of Bacon County, where there was a need to create new ballots for the presidential race to replace those that were damaged or unreadable from the initial count. Notably, Bartow County demonstrated an approach in their duplication process, ensuring that the proper combo codes and districts were used for the recreated ballots.

Investigation into Specific Ballots

The aspect that draws attention in Bartow County, and forms the crux of this analysis, is the nature and source of the ballots that required duplication. Unlike Bacon County, where the 23 source ballots needing duplication were diverse and from multiple batches, Bartow County's scenario was markedly different:

Specific Tabulator and Batches Involved

- Tabulator 530 – Batch 185: 36 sequential ballots
- Tabulator 530 – Batch 189: 11 ballots
- Tabulator 530 – Batch 190: 46 sequential ballots

Points of Curiosity and Speculation

The pattern observed in Bartow County raises certain questions:

- Sequential Nature of Ballots: The sequential nature of the ballots, particularly in Batches 185 and 190, is intriguing. This suggests a systematic issue rather than random occurrences, which is more common in such scenarios.
- Concentration in Specific Batches: The fact that the ballots needing recreation came exclusively from the same tabulator, and specifically from two entire batches, suggests a potential external factor influencing these batches. Could there have been environmental factors, such as water damage, that led to the failure of these entire batches?

- Comparison with Other Counties: In contrast to other counties where duplicate ballots are generally random and attributed to printing issues, Bartow's case appears unique. The need to recreate two full batches of ballots is unusual and warrants further investigation to understand the underlying cause.

Conclusion

Bartow County's case stands out in the ballot analysis for its specific challenges and the efficient response in duplicating the affected ballots. However, the peculiar pattern observed in the duplication process invites curiosity and speculation about the possible causes. Understanding the factors leading to such a concentrated need for ballot duplication in specific batches can provide valuable insights into the robustness of the voting process and potential areas for improvement in ballot handling and preservation.

Bibb County

For Bibb County, the source of the information relies on images and data files from the following sources.

- MC1 TIF Ballot Images (Incomplete)
- MC2 TIF Ballot Images (70,917 records)
- MC1 CSV Cast Vote Record (71,170 records)
- MC2 CSV Cast Vote Record (71,088 records)

Challenges in Analysis

Missing BMD Ballots: The original batch of ballots lacked some Ballot Marking Device (BMD) ballots, complicating the analysis.

Conversion of Ballots: A significant number of ballots were converted from Hand Marked Paper Ballots (HMPB) to BMD ballots. While these converted ballots were properly marked and highly legible, their interpretation posed challenges.

CVR Misreading: The converted ballots, primarily from absentee and provisional categories, were initially misinterpreted in the Cast Vote Record (CVR). This misinterpretation led to discrepancies in the ballot signatures, which were later rectified by adjusting the initial characters in the signatures. The total number of duplicated ballots in the original count was found to be 402.

Discrepancies Between Counts

In the recount, 103 ballots from the original count were absent. Out of these, five were HMPB and the remaining were BMD ballots.

Distribution of Stray Ballots: Of the missing ballots, 75 were votes for Joe Biden and 27 for Donald Trump.

Additional Ballots in Recount

In the MC2 recount, there were 22 ballots that were not present in the original MC1 count. Among these, five were BMD ballots and the rest HMPB.

Voting Pattern of New Ballots: Of these ballots, 19 votes were for Joe Biden and 4 for Donald Trump.

Conclusion

The ballot analysis for Bibb County highlighted several challenges, particularly in reconciling discrepancies between different ballot types and counts. The data indicates a mix of missing and additional ballots across the original count and the recount, with variations in the distribution of votes for the presidential candidates.

Catoosa County

For Catoosa County, the source of the information relies on images and data files from the following sources.

- MC1 TIF ballot images (32,725 images)
- MC2 TIF ballot images (32,756 images)

Overview of Ballot Counts

In Catoosa County, the analysis of ballot data presented a relatively straightforward scenario. However, the recount showed a slight increase, which marked a difference of 31 ballots.

Analysis of Discrepancies

New Ballots in Recount: A total of 29 ballots were identified in the recount that were not present in the original count. The breakdown of these additional ballots is as follows:

- 4 ballots from Tabulator 310, Batch 1.
- 6 ballots from Tabulator 320, Batch 1.
- 18 ballots from Tabulator 320, Batch 2.

Notable Observations

Concentration on Tabulator 320: A significant observation was that the majority of the new ballots appeared on Tabulator 320, which processed only two batches.

Voting Pattern of New Ballots: These additional ballots predominantly favored Joe Biden.

Order of Ballot Scanning: Contrary to patterns observed in other counties where new ballots typically appeared at the end of the count, in Catoosa County, these ballots from Tabulator 320 were the first to be scanned during the recount.

Conclusion

The analysis of Catoosa County's ballots showed a relatively small discrepancy between the original count and the recount. The additional ballots found in the recount were concentrated in specific batches and tabulators, notably favoring one presidential candidate.

Cherokee County

For Cherokee County, the source of the information relies on images and data files from the following sources.

- MC1 TIF ballot images (145,564 images)
- MC2 TIF ballot images (145,545 images)
- MC1 CSV Cast Vote Record (145,564 records)
- MC2 CSV Cast Vote Record (145,545 records)

Key Findings

Double-Scanned Ballots: The original count included 125 double-scanned ballots, primarily originating from Tabulators 950 and 955. Notably, these tabulators were responsible for duplicating three batches from Tabulator 950.

Stray Ballots in Original Count: There were 22 stray ballots identified exclusively in the original count. The majority of these stray ballots were Ballot Marking Device (BMD) ballots from the Hillside precinct.

Stray Ballots in Recount: Interestingly, the recount presented 127 stray ballots, a figure closely resembling the number of double-scanned ballots identified in the original count. These additional ballots were mostly Hand Marked Paper Ballots (HMPB) from the Air Acres and Woodstock precincts.

Unique Case in Rosecreek Precinct

A peculiar situation was observed in the Rosecreek precinct, where 22 new ballots surfaced in the recount, featuring selections only in the U.S. Senate race. These ballots notably lacked any entries for the presidential race and other down-ballot contests. In the context of ballot signatures, the term 'BLA' is used to denote a blank entry in a contest.

Observations and Questions

- The appearance of 20 ballots in the recount, marked solely for the senatorial race, raises questions. The reason behind the creation of these specific ballots for the recount remains unclear, warranting further investigation to understand the motive or procedural context that led to this anomaly.

Conclusion

The ballot analysis in Cherokee County underscores the need for attention to detail in the electoral process. While the overall discrepancy in ballot counts was minimal, the presence of double-scanned and stray ballots, as well as the unique case in the Rosecreek precinct, highlight the complexities involved in ensuring accuracy and integrity in vote tabulation.

Attached is a copy of one of the ‘senator only’ ballots

CHEROKEE COUNTY

OFFICIAL BALLOT

GENERAL AND SPECIAL ELECTION
OF THE STATE OF GEORGIA
NOVEMBER 3, 2020

"I understand that the offer or acceptance of money or any other object of value to vote for any particular candidate, list of candidates, issue, or list of issues included in this election constitutes an act of voter fraud and is a felony under Georgia law." [O.C.G.A. 21-2-284(e), 21-2-285(h) and 21-2-383(a)]

213-Rosecreek

For President of the United States (Vote for One) (NP)
BLANK CONTEST

For United States Senate (Perdue) (Vote for One) (NP)
BLANK CONTEST

For United States Senate (Loeffler) - Special (Vote for One) (NP)
Vote for Derrick E. Grayson (Rep)

For Public Service Commissioner (Vote for One) (NP)
BLANK CONTEST

For State Representative in the General Assembly From 20th District (Vote for One) (NP)
BLANK CONTEST

For District Attorney of the Blue Ridge Judicial Circuit (Vote for One) (NP)
BLANK CONTEST

For Judge of the Probate Court (Vote for One) (NP)
BLANK CONTEST

For Clerk of Superior Court (Vote for One) (NP)
BLANK CONTEST

For Chief Magistrate (Vote for One) (NP)
BLANK CONTEST

For County Board of Education Chairman (Vote for One) (NP)
BLANK CONTEST

For Limestone Valley Soil and Water Conservation District Supervisor (Vote fo One) (NP)
BLANK CONTEST

Constitutional Amendment #1 (NP)
BLANK CONTEST

Constitutional Amendment #2 (NP)

Cobb County

For Cobb County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (396,551 records)
- MC2 CSV CVR files (396,517 records)
- MC2 TIF ballot images (396,532 images)

Overview of Ballot Counts

In Cobb County, the analysis of ballot data presents a scenario with notable discrepancies. The recount (MC2) showed a slightly lower count indicating a loss of 34 ballots. Additionally, the MC2 ballot images contained 15 more records than the Cast Vote Record.

Challenges in Analysis

Lack of Images for First Count: A significant hindrance in the analysis process is the absence of images from the first machine count. This limitation makes it challenging to conduct a conclusive analysis of the county. Once the original ballot images are obtained, it is anticipated that several hundred more ballot issues may be identified, particularly related to signature mismatches.

Discrepancies Identified

Ballots Missing in Recount: There were at least 481 ballots present in the first count that did not appear in the recount.

Duplicate Ballots in Recount: In the recount, there are indications of at least 428 ballots being duplicated or not found in the original count:

- Approximately 400 ballots were identified as double-scanned from various previous batches.
- 28 stray ballots were found in the recount that were not part of the original count.

Issues with Duplicate Ballots

Random Order of Duplicates: Some of the duplicate ballots in the recount appeared to be quite mixed up and in a random order compared to their arrangement in the original count.

Verification Challenges: The lack of actual ballot images from Cobb County's first count complicates the verification process of the exact ballots that were duplicated. However, the current analysis is believed to be very close to accurate, pending the receipt of the original ballot images.

Conclusion

The ballot analysis for Cobb County highlights the importance of having complete data sets, including original ballot images, for an accurate and thorough analysis.

DeKalb County

For DeKalb County, the source of the information relies on images and data files from the following sources.

- MC1 TIF ballot images (373,439 images)
- MC2 TIF ballot images (373,367 images)

Overview of Ballot Counts

DeKalb County's ballot analysis indicated a small discrepancy between the two counts resulting in a difference of 72 records.

Discrepancies and Findings

Missing Ballots in Second Count: A total of 129 ballots were present in the original count but not included in the second count. Remarkably, all but four of these were Ballot Marking Device (BMD) ballots. The missing ballots were exclusively from Tabulator 5040, primarily involving the removal of four batches.

New Ballots in Recount: Conversely, the recount added 160 ballots that were not part of the original count.

Specific Cases of Interest

Addition of Ballots: A noteworthy addition of 105 ballots was observed in Tabulator 5050, Batch 145. These ballots were distinctly marked with the word 'Recount.' This scenario mirrors patterns seen in Bacon and Bartow counties, where new presidential ballots were presumably created as duplicates for damaged ballots.

Nature of Adjusted Ballots: Most of these adjusted ballots were limited to the presidential vote, indicated by the signature 'BLA' for blank entries. The coding 'Don' and 'Jos' were used to represent votes for Donald Trump and Joseph Biden, respectively.

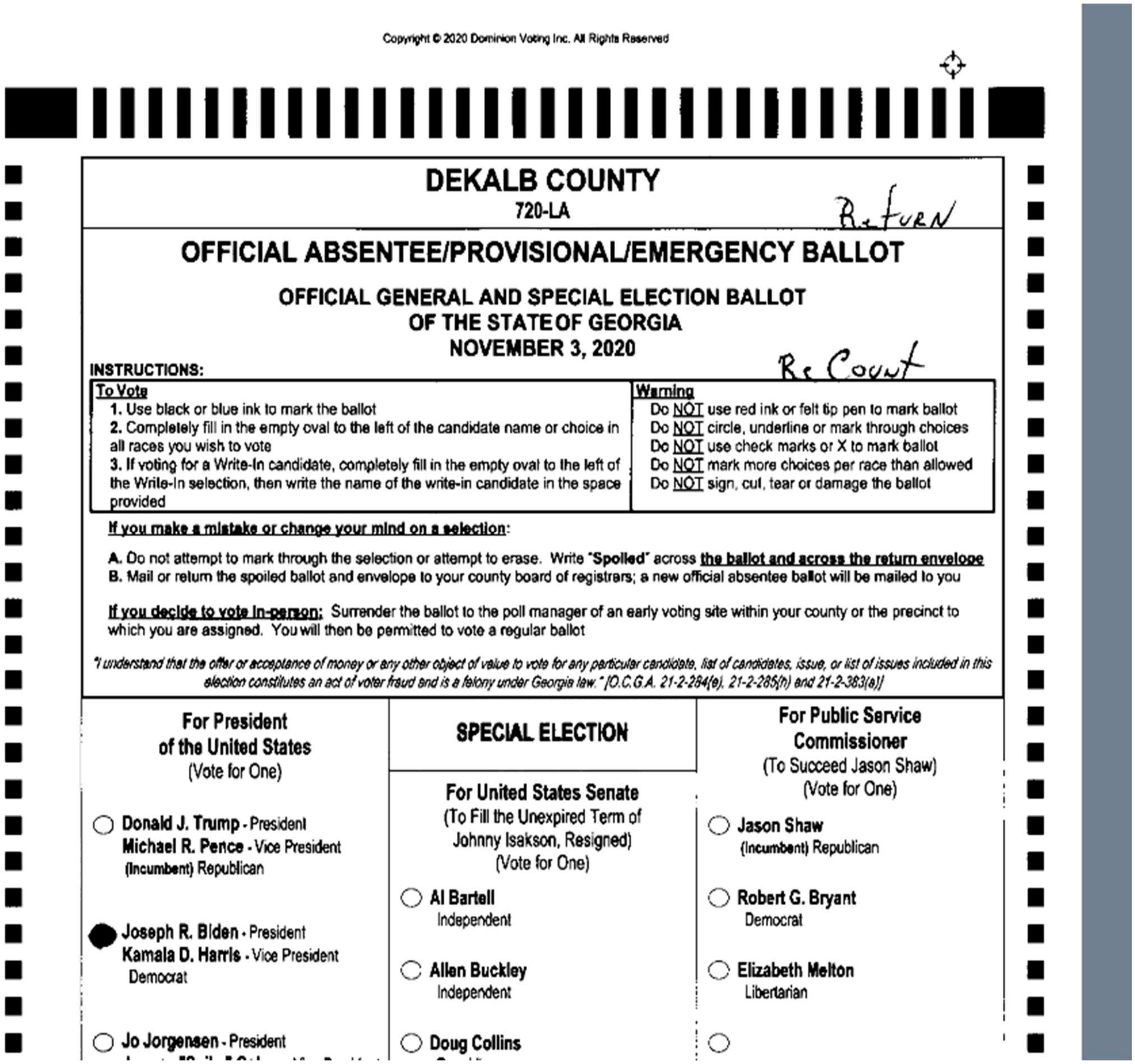
Labeling for Clarity: The labeling of these ballots with 'Recount' by a diligent worker played a crucial role in identifying these unusual additions, providing clarity and transparency in the recount process.

Recreation of Whole Batches: The need to recreate entire batches as 'recount' ballots raises questions. The rationale behind replacing a significant number of ballots from the original count with newly created 'recount' ballots remains unclear.

Conclusion

Despite these peculiarities, it is important to acknowledge the overall efficiency of DeKalb County in the ballot counting process. Considering the challenges posed by a large population, the county performed commendably in managing and executing the count.

The ballot analysis for DeKalb County highlights a few instances of irregularities, particularly in the creation and inclusion of new ballots marked for recount. These anomalies, while small in number compared to the total count, underscore the importance of transparent and consistent procedures in ballot handling and tabulation.



Copyright © 2020 Dominion Voting Inc. All Rights Reserved

DEKALB COUNTY
720-LA

ReturN

OFFICIAL ABSENTEE/PROVISIONAL/EMERGENCY BALLOT

OFFICIAL GENERAL AND SPECIAL ELECTION BALLOT
OF THE STATE OF GEORGIA
NOVEMBER 3, 2020

Re Count

INSTRUCTIONS:

To Vote	Warning
1. Use black or blue ink to mark the ballot 2. Completely fill in the empty oval to the left of the candidate name or choice in all races you wish to vote 3. If voting for a Write-In candidate, completely fill in the empty oval to the left of the Write-In selection, then write the name of the write-in candidate in the space provided	Do NOT use red ink or felt tip pen to mark ballot Do NOT circle, underline or mark through choices Do NOT use check marks or X to mark ballot Do NOT mark more choices per race than allowed Do NOT sign, cut, tear or damage the ballot

If you make a mistake or change your mind on a selection:

A. Do not attempt to mark through the selection or attempt to erase. Write "Spoiled" across **the ballot and across the return envelope**
B. Mail or return the spoiled ballot and envelope to your county board of registrars; a new official absentee ballot will be mailed to you

If you decide to vote in-person: Surrender the ballot to the poll manager of an early voting site within your county or the precinct to which you are assigned. You will then be permitted to vote a regular ballot

"I understand that the offer or acceptance of money or any other object of value to vote for any particular candidate, list of candidates, issue, or list of issues included in this election constitutes an act of voter fraud and is a felony under Georgia law." [O.C.G.A. 21-2-284(e), 21-2-285(h) and 21-2-383(a)]

For President of the United States
(Vote for One)

- Donald J. Trump - President
  Michael R. Pence - Vice President
  (Incumbent) Republican
- Joseph R. Biden - President
  Kamala D. Harris - Vice President
  Democrat
- Jo Jorgensen - President

SPECIAL ELECTION

For United States Senate
(To Fill the Unexpired Term of Johnny Isakson, Resigned)
(Vote for One)

- Al Bartell
  Independent
- Allen Buckley
  Independent
- Doug Collins

For Public Service Commissioner
(To Succeed Jason Shaw)
(Vote for One)

- Jason Shaw
  (Incumbent) Republican
- Robert G. Bryant
  Democrat
- Elizabeth Melton
  Libertarian

Floyd County

For Floyd County, the source of the information relies on images and data files from the following sources.

- MC1 TIF ballot images (38,666 records)
- MC2 TIF ballot images (41,573 records)

Floyd County, Georgia, found itself at the center of attention during a recount in the 2020 presidential race. This recount unveiled a significant discrepancy in the vote tally, highlighting the challenges and the importance of accuracy in the electoral process.

As reported in an article by The Atlanta Journal-Constitution, the recount in Floyd County unearthed over 2,600 ballots that had not been included in the original count. This discovery had the potential to slightly narrow President Donald Trump's 14,000-vote deficit to Joe Biden in the state. The breakdown of these newfound votes was particularly noteworthy: Trump received an additional 1,643 votes, Biden 865, and Jo Jorgensen with 9.

The root cause of this discrepancy was attributed to human error rather than a technical malfunction. The issue stemmed from a failure to upload votes from a memory card in a ballot scanning machine. The uncounted votes predominantly came from in-person early voting conducted at the Floyd County Administration Building, which houses the county's elections office. It was reported that more than half of the 5,000 ballots cast using an optical scanner at this location were not initially recorded.

Revised Analysis of Ballot Counts in Floyd County

Initial assessments suggested that the count discrepancy in Floyd County involved approximately 2,600 ballots. However, a closer examination of the ballot numbers revealed a more significant variance. The count from Machine Count 1 (MC1) to MC2 showed a disparity of 2,907 ballots – a difference notably larger by 323 ballots than initially estimated.

Utilizing the count vs. recount detection feature of my software, I identified all new ballots that emerged in the recount. The comprehensive analysis yielded the following findings:

In Machine Count 1 (MC1):

- 4 ballots were identified as double-scanned.
- 12 stray ballots were present in MC1 but not found in MC2.

In Machine Count 2 (MC2):

- A substantial addition of 2,923 ballots that were not included in MC1.

Distribution of the Missing Votes Among Presidential Candidates:

- Donald Trump received an additional 1,904 votes.
- Joe Biden received 975 additional votes.
- Jo Jorgensen garnered an extra 22 votes.

These findings significantly alter the understanding of the vote totals in Floyd County, highlighting the importance of thorough ballot analysis to ensure accuracy in election results.

Forsyth County

For Forsyth County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (129,945 records)
- MC2 CSV CVR files (129,947 records)
- MC2 TIF ballot images (129,970 records)

Ballot Counts Overview

In Forsyth County, the ballot count analysis showed a close alignment between the different sources, with minor discrepancies. The cast vote records recorded between 129,945 to 129,947 ballots. However, the number of ballots in the TIF images was 25 records higher.

Primary Issue Identified

Double-Counted Records: The main issue observed in Forsyth County was a specific instance of double counting. This occurred in a batch from Tabulator 653, Batch 246, where 23 records were double-counted, inadvertently rescanning the previous Batch 245.

Confirmation in MC2: It was confirmed that these double-counted ballots were included in the MC2 Cast Vote Record, contributing to the count discrepancy.

Discrepancy between CSV and TIF Images

Despite the identification of the double-counted ballots, a lingering question remains regarding the discrepancy of 23 votes between the CSV records and the TIF images. The reason behind this variance is not immediately clear and warrants further investigation.

Limitations in Analysis

Limited Scope of Cast Vote Record: A significant limitation in the analysis for Forsyth County was that the provided Cast Vote Record only included data for the presidential ballot. This restriction made it challenging to conduct a comprehensive ballot analysis, as it limited the scope of data available for review and comparison.

Conclusion

While the overall discrepancy in ballot counts was relatively small, the issue of double-counted records and the unexplained discrepancy between CSV and TIF images underscore the need for careful management and verification of voting records. Additionally, the limited scope of the provided Cast Vote Record for only the presidential ballot presents a challenge in conducting thorough and multifaceted ballot analysis.

Fulton County

For Fulton County, the source of the information relies on images and data files from the following sources.

- MC1 ballot images (152,209 images)
- MC2 ballot images (510,202 images)
- MC1 JSON CVR files (528,776 records)
- MC2 JSON CVR files (527,926 records)

Overview

Fulton County presented significant challenges in the analysis of ballot data, with an unusually high number of discrepancies observed. The complexity and volume of issues in Fulton County were remarkably higher than those found in any other county examined.

Discrepancies in Ballot Counts

The analysis revealed notable discrepancies between the original count and the recount in terms of duplicated and stray ballots.

Double/Triple Scanned Ballots

- In the original count, there were 391 doubled ballots.
- The recount showed a substantially higher number of doubled/tripled ballots, totaling 3,920.

Stray Ballots

- A total of 3,284 stray ballots were identified in the original count but not in the recount.
- Conversely, the recount included 1,413 stray ballots that were not present in the original count.

Challenges in Consolidation

The process of consolidating and analyzing the ballots for Fulton County was particularly daunting due to several factors:

- Volume of Issues: The sheer number of mismatched ballots in Fulton County surpassed the total number found in all other counties combined, indicating a high level of complexity in the voting records.
- Accuracy of Double-Scanned Ballot Count: The count of double-scanned ballots is believed to be very precise, especially since most of the doubled Hand Marked Paper Ballot (HMPB) images were available for examination.
- Ongoing Stray Ballot Analysis: The assessment of stray ballots is still a work in progress, with further study and interpretation ongoing. Therefore, the current figures for stray ballots are subject to change as the investigation continues.

Additional Complications

- Missing Source Ballot Images: Adding to the complexity, most of the source ballot images from the first Machine Count were missing. This absence has significantly hindered the process of accurately determining stray versus duplicate ballots.
- Extended Duration of Analysis: Efforts to resolve the aberrations in Fulton County have been ongoing for several months with limited progress. The extensive number of issues to be addressed in this county presents a formidable challenge, far exceeding the scope of work required for other counties.

Grady County

For Grady County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (10,741 records)
- MC2 CSV CVR files (10,740 records)
- MC1 TIF ballot images (10,790 images)
- MC2 TIF ballot images (10,740 images)

Key Finding

Double-Scanned Batch: The primary issue in Grady County involved a batch of 49 ballots that were double-scanned. This occurred in Tabulator 290, Batch 48, which inadvertently duplicated Batch 1. This instance of double scanning contributed to the discrepancy in the TIF image counts.

Discrepancy Between TIF Images and CSV

CVR Count Anomaly: Despite the lower count in the original CVR, suggesting that duplicate ballots might have been removed, it was found that these double-scanned ballots were still present in the CSV data. This implies that another batch of ballots was omitted in the CSV, but the specific batch and the reasons for its removal remain unclear.

Pending Investigation

Unidentified Missing Batch in CSV: There is a need for further investigation to identify the missing batch in the CSV data. Understanding which ballots were removed and why will be crucial in reconciling the discrepancies between the different data sources and ensuring the accuracy of the final vote tally.

Conclusion

The analysis of Grady County's ballot data underscores the challenges in maintaining consistency across different formats of vote tabulation.

Gwinnett County

For Gwinnett County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (536,616 records)
- MC2 CSV CVR files (536,654 records)
- MC1 JSON CVR files
- MC1 TIF ballot images (536,616 images)
- MC2 TIF ballot images (536,226 images)

Context of Ballot Processing

Gwinnett County presented a unique challenge in ballot processing due to the use of two ballots per voter. A secondary constitutional ballot page was included, which was later separated from the primary ballot during the hand count. This separation added complexity to the analysis process.

Key Findings

Doubled Ballots in First Count: There were 449 ballots identified as doubled in the first count, originating from nine different batches.

Stray Ballots Analysis: Investigation into stray ballots that were present in the original count but missing in the recount identified 575 such cases.

Ballots Missing in Recount: A total of 963 ballots present in the original count were not found in the recount.

Double Scanned Ballots in Recount: The recount also revealed 151 double-scanned ballots.

More records than images: There are 428 more Cast Vote Records in MC2 than ballot images.

Recount Specifics

Removal of Non-Presidential Votes: During the recount, which focused primarily on the presidential vote, most ballots without a presidential entry were removed from the count. This decision led to a large number of stray ballots in the original count, though it did not affect the presidential vote tally.

Issue with Constitutional Ballots

Separation of Constitutional Ballots: During the recount, constitutional ballots were separated from the candidate ballots, making it challenging to identify stray constitutional ballots. It is estimated that approximately 250 additional stray constitutional ballots were not included in the recount count.

Presidential Vote Proportion: Given the dual ballot system in Gwinnett County, it's important to note that roughly half of the votes were for a presidential candidate. This aspect is crucial in understanding

the distribution and significance of the votes and any discrepancies that arose during the counting process.

Conclusion

The ballot analysis for Gwinnett County underscores the complexities introduced by using two ballots per voter and the subsequent separation of these ballots during the recount. The reduction in the ballot count in the recount, coupled with issues like ballot duplication, stray ballots, and the challenge in tracking constitutional ballots, highlights the need for a careful and thorough examination of the voting and recount processes

Houston County

For Houston County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (75,169 records)
- MC1 TIF ballot images (75,166 images)
- MC2 TIF ballot images (75,187 images)

Overview of Discrepancies

In Houston County, the ballot counts between the initial count and the recount revealed notable discrepancies, albeit on a smaller scale compared to some other counties.

Missing Ballots in Second Count: There were four ballots from the initial count that were not accounted for in the recount. This included two Hand Marked Paper Ballots (HMPB) and two Ballot Marking Device (BMD) ballots.

New Ballots in Recount: Conversely, the recount showed the appearance of 28 new ballots, the majority of which were BMD ballots. These ballots did not follow any discernible pattern and were found across various batches.

Peculiarities in Ballot Placement

Scattered Placement of New Ballots: One unusual aspect of these additional ballots was their scattered placement at the end of several batches. This irregularity is atypical, as new or doubled ballots are generally grouped together, suggesting a more systematic approach to their inclusion.

Uncommon Duplication of BMD Ballots: The duplication of BMD ballots in this manner is uncommon. BMD ballots are typically more controlled and less prone to duplication compared to HMPB ballots, given their automated and standardized nature.

Conclusion

The ballot analysis for Houston County points to peculiarities that warrant further investigation, particularly regarding the scattered placement and the nature of the new ballots in the recount. The lack of a clear, organized method in the inclusion of these additional ballots, especially the BMD ballots, raises questions about the ballot handling and tabulation processes.

Muscogee County

For Muscogee County, the source of the information relies on images and data files from the following sources.

- MC1 ballot images (80,977 images)
- MC2 ballot images (81,346 images)

Overview of Doubled Ballots in First Count

In Muscogee County, the analysis revealed instances of ballot duplication in the first count:

- 25 ballots from Tabulator 550, Batch 95 were found to be doubled of those in Batch 92.
- Similarly, 25 ballots from Tabulator 550, Batch 567 doubled those in Batch 566.

Discrepancies Between Counts

The recount showed a significant discrepancy with 434 ballots from the original count missing. This included:

- 270 BMD ballots from Tabulator 273.
- 53 BMD ballots from Tabulator 274.
- 56 HMPB ballots from Tabulator 550.

Puzzling Aspects

BMD Ballot Discrepancy: The disappearance of 378 BMD ballots in the recount is highly unusual, especially considering many were write-in ballots from Tabulator 273. The likelihood of these ballots being separated during the manual recount, while plausible, is uncommon and necessitates further examination.

Discovery of New Ballots in Recount

The recount surprisingly revealed 440 new ballots that were not included in the original count:

- 208 HMPB ballots from Tabulator 650.
- 230 HMPB ballots from Tabulator 750.

Doubled Ballots in Second Count

The second count identified 455 doubled ballots. The nature of these doubles is intriguing, with mismatches across six different batches. Some batches contributed individual ballots, while others had multiple ballots doubled. Adding to the complexity, these ballots were scanned in reverse order, likely due to being placed upside down in the scanner.

Conclusion

The ballot analysis for Muscogee County uncovers a complex scenario with multiple layers of irregularities. The disappearance of a substantial number of ballots in the recount, coupled with the discovery of new ballots and the peculiar pattern of duplications, raises several questions about the ballot handling and counting processes.

Upson County

For Upson County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (12,951 records)
- MC2 CSV CVR files (12,950 records)
- MC2 TIF ballot images

Overview of Ballot Counts

The mere discrepancy of one vote between the two counts is noteworthy.

Investigating the Narrow Discrepancy

- Stray Ballots in Recount: Despite the minimal difference in overall ballot counts, there appears to be a significant number of missing BMD ballots – over 101 – missing from the recount. Under normal circumstances, such a number of stray ballots would have led to a more substantial discrepancy in the total count.
- Presence of Double Scanned Ballots: Intriguingly, the presence of approximately 100 double scanned ballots seems to counterbalance the missing stray ballots. This coincidence of numbers is unusual and effectively resulted in aligning the vote counts between the original tally and the recount.

The Need for a Closer Look

The striking balance achieved in Upson County between stray and double scanned ballots raises questions about the counting process and the accuracy of the results. The precision with which these two factors offset each other to produce nearly identical total counts is a phenomenon that warrants a more detailed investigation.

Upson County presents another case where the interplay of stray and double scanned ballots has led to an unexpectedly close alignment in vote counts between the original count and the recount. The apparent accuracy of the final tally, despite the presence of discrepancies in ballot processing, indicates a need for a deeper dive into the county's counting methods and practices to fully understand how such a result was achieved.

Walton County

For Walton County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (51,010 records)
- MC2 CSV CVR files (51,289 records)
- MC2 TIF ballot images

Adjudication Findings

Added Ballots in Recount: The adjudication process in Walton County revealed that 284 additional ballots were incorporated into the recount. This number slightly exceeds the initial discrepancy of 279 ballots identified between the two counts.

Distribution of Additional Votes

Of these 284 new ballots, 225 votes were for Donald Trump, 49 for Joseph Biden, and 10 for Jo Jorgensen.

Source of New Ballots: It was determined that all the new ballots originated from Ballot Marking Devices (BMD) machines.

Conclusion

The analysis of Walton County's ballots indicates a fairly straightforward process with a clear addition of ballots in the recount. The majority of these additional votes favored one presidential candidate. The exclusive origin of these new ballots from BMD machines highlights the role of these devices in the voting process and the importance of their accurate functioning and monitoring.

Ware County

For Ware County, the source of the information relies on images and data files from the following sources.

- MC1 CSV CVR files (14,404 records)
- MC1 TIF ballot images (14,307 records)
- MC2 TIF ballot images (14,253 records)

Adjudication Findings

There were 100 ballots that were doubled in the first count.

- 100 ballots from Tabulator 270, Batch 35 matching Tabulator 270, Batch 11
- 2 ballots from Tabulator 270, Batch 32 that were only in the first count.

In the second count, 48 new were found

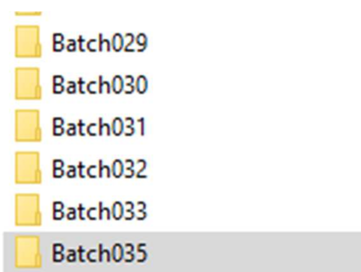
- 48 ballots from Tabulator 270, Batch 29

The increase of 97 ballots in the cast vote record for the recount has not yet been identified.

Pending Investigation

The situation with Ware County presents a curious case of irregularity. The fact that the 100 doubled ballots were all processed at the very end of the initial count raises eyebrows.

Compounding this mystery is the perplexing absence of Batch 34 from the results. Its conspicuous absence raises a host of questions: Was it accidentally omitted, or was it deliberately excluded?



The unexplained appearance of new ballots in the recount adds another layer of complexity to the situation. The seemingly arbitrary addition of 50 ballots to the last batch raises the possibility that these could be the missing Batch 34.

The last observation is the most curious. The count of the ballots was completed between 11/2/2020 and 11/6/2020.

The doubles were added on 11/11/2020. 75 votes for Biden, 25 for Trump.

Database Security Deficiencies

Introduction

In a world increasingly reliant on digital systems, the security of electoral data stands paramount. This chapter delves into concerning revelations about the security measures—or the lack thereof—surrounding voter databases in several counties in Georgia. These findings came to light following open record requests made by VoterGA, a group dedicated to ensuring the integrity and transparency of the voting process.

Accompanying the requested ballot information, backups of voter databases from four counties—Appling, Bibb, Jones, and Telfair—were included. Astonishingly, these databases could be accessed effortlessly using Microsoft's SQL Server Management Studio (SSMS), a standard tool for managing databases. The absence of security layers or password protection to access these databases was the first red flag, indicating a fundamental oversight in protecting sensitive electoral data.

Upon exploring the contents of these databases, I made several observations that further underscore the pressing need for robust security measures in our electoral systems. The lack of basic security protocols, especially concerning password management and data encryption, was alarming. These vulnerabilities not only raise questions about the potential for unauthorized access but also bring to light broader issues of data integrity and trust in our electoral infrastructure.

In this chapter, I will document these security deficiencies in detail, drawing from the insights gained through the analysis of these voter databases. The aim is to highlight the critical importance of cybersecurity in the electoral process and to call for immediate action to safeguard our democratic institutions from digital threats.

Technician & Supervisor Tabulator Passwords

Upon successfully accessing the database, my attention turned to a particularly intriguing table named **`SystemParameter`**. This table appeared to be a repository of various system settings and configurations, an essential aspect of any database. To gain insights specifically into the security protocols, I formulated and executed the following SQL query:

```
SELECT paramName, paramValue, description
FROM SystemParameter
WHERE description LIKE 'passcode for%'
```

The results from this query were both surprising and alarming:

paramName	paramValue	description
Technicians Passcode	**123456**	Passcode for technicians. Used for rezeroing and reopening the poll.
Technicians Passcode	**870913**	Passcode for technicians. Used for rezeroing and reopening the poll.
Supervisor Passcode	**123456**	Passcode for supervisors. Used for rezeroing and reopening the poll.

What stood out immediately was the revelation of plaintext passwords within the database, a cardinal sin in modern cybersecurity practices. These passwords, meant for technicians and supervisors to rezero and reopen polls, were not only visible without any encryption but also, in many instances, alarmingly simplistic.

The use of **`123456`** as a default password is a glaring example of inadequate security measures. Such passwords are notoriously weak and easily guessable, making them vulnerable to even the most basic forms of unauthorized access.

Interestingly, the table contained two entries for each county with varying values. Along with the standard **`123456`** password (which incidentally is the same password on my luggage), in Bibb County, the password for technicians was set to **`870913`**.

The ease with which these passwords were accessed and the apparent lack of complexity in password selection raise significant concerns about the overall security posture of these systems. The fact that such rudimentary and easily exploitable practices were in place within a voting system's infrastructure is not just concerning; it borders on negligence. This section aims to delve deeper into these security flaws and their potential implications on the integrity of the electoral process.

The technician and supervisor passwords for the various counties are

- Appling: **`515250`**
- Bibb: **`870913`**
- Jones: **`356753`**
- Telfair: **`228823`**

Uncovering the **`ElectionEvent`** Table

Progressing further into the database, my investigation led me to a table named **`ElectionEvent`**, a repository of ostensibly benign information related to various electoral events. This table included standard fields like the date of the election and descriptive labels, such as 'Bibb County November 2020 General and Special Election'. However, it was the discovery of certain other fields in this table that raised immediate and profound security concerns.

The Concerning Fields:

1. **RijndaelKey** and **RijndaelVector**: These fields are particularly alarming. The Rijndael algorithm is a form of encryption, and these fields — the Key and Vector — are critical components of the encryption process. For example, the Rijndael Key was listed as **`Ka1&!Bt5f9J?Me%2`**, and the Rijndael Vector as **`6El%]1Dgt^7R$8Si`**. The unencrypted visibility of these values is a significant lapse in security protocol. With access to these, one could potentially decrypt passwords and sensitive election files stored as "DVD" files in the system.

2. **X509Data**: This field contained a byte array representing a full x.509 security certificate, including the crucial public/private key pair. X.509 certificates are a standard for public key infrastructure and are used for secure data exchange, including digital signatures. The exposure of this certificate, especially the private key, is a major vulnerability, effectively laying bare the system's secure communication channels.

3. **HMACKey**: Although its specific application in this context was unclear, HMAC (Hash-based Message Authentication Code) Keys are generally used for data integrity checks and authentication. The visibility of this key further adds to the list of security lapses.

Name	RjindaelKey	RjindaelVector	X509Data	HMAC
Bibb Nov 2020 General	Ka1&!Bt5f9J?Me%2	6El%]1Dgt^7R$8Si	0x308205E	0x326B7C

Implications of These Findings:

The presence of these unencrypted keys in the database is a glaring security oversight. It suggests that the electoral systems might be vulnerable to unauthorized decryption and data manipulation. Particularly concerning is the ease with which I could potentially use the Rijndael keys and vectors, even via a public website, to decrypt user passwords.

In subsequent chapters, I will delve deeper into the implications of these discoveries, particularly how the exposed x.509 certificate compromises the system's integrity. The findings underscore a pressing need for a comprehensive review and overhaul of the security measures protecting these critical electoral systems.

Exploring the **`TabulatorUser`** Table

Continuing my investigation into the security protocols of the electoral system, I turned my attention to a database table named **`TabulatorUser`**. This table ostensibly listed credentials for tabulator administrators, and my findings here were both revealing and concerning.

The Surprising Discovery:

Upon querying the table, the results were quite unexpected. The table comprised 89 rows, but strikingly, each row was identical:

username	password	firstName	lastName	type
Admin	NLNA2TeVbYW4eAm58bDDfA==	Admin	Admin	Poll Worker
Admin	NLNA2TeVbYW4eAm58bDDfA==	Admin	Admin	Poll Worker
...	...	...	...	...

The data showed that all 89 Bibb county poll workers, along with those in other counties like Appling (15 rows), Bibb (20 rows), and Telfair (14 rows), shared the same password and access levels.

Decryption of the Passwords:

Although the passwords were not in plaintext, each county had its distinct encrypted values. Utilizing the Rijndael Key and IV from each County, I tested their decryption on a public website.

- Telfair County: **71440304**
- Bibb County: **54762176**
- Jones County: **76365112**
- Appling County's **61837066**

The simplicity of this decryption process was alarming. It required minimal effort and basic encryption knowledge, suggesting that anyone with such skills could potentially access these passwords.

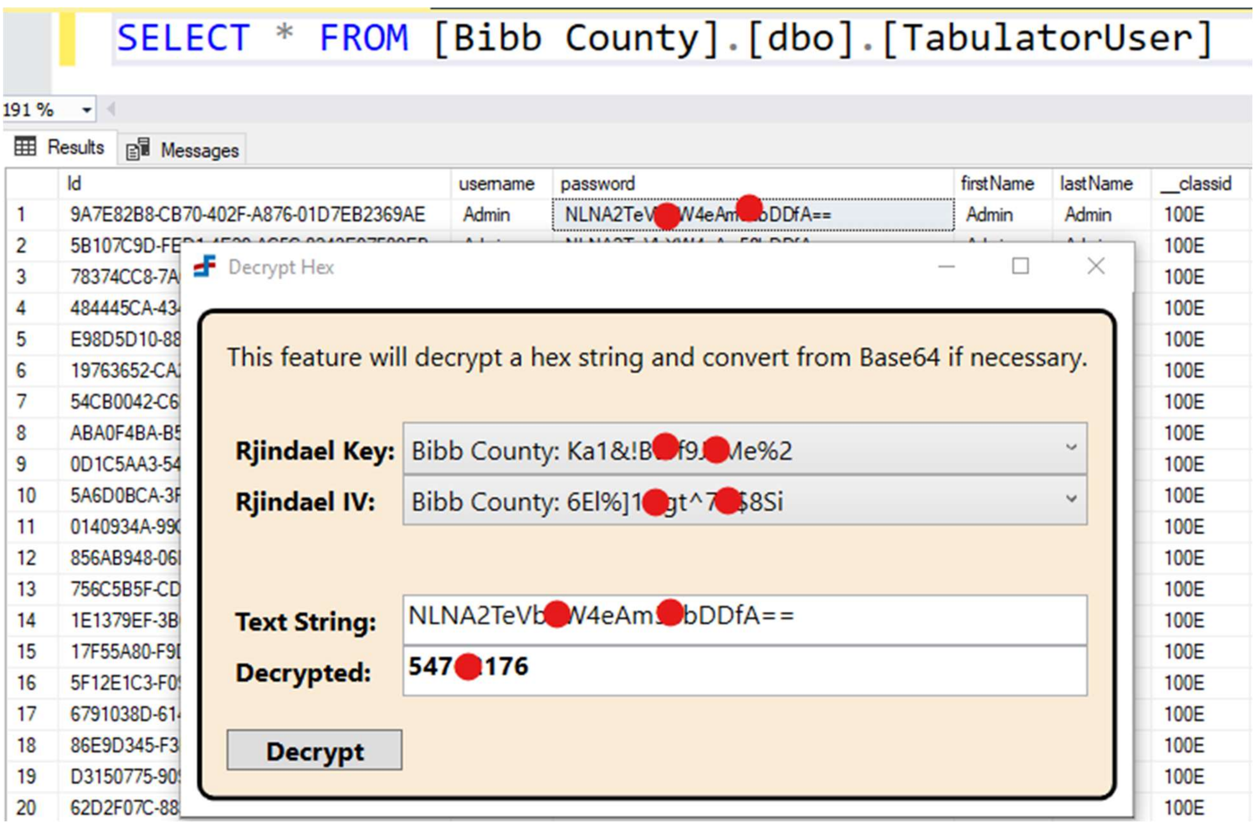
Critical Security Lapses:

Two major security lapses were evident from these findings:

Improper Password Storage: In the realm of cybersecurity, it is a standard practice to hash passwords using a one-way hash rather than encrypt them. This method ensures that even if someone gains access to the password hash, they cannot decrypt it to obtain the original password. The electoral system's use of reversible encryption rather than hashing demonstrated a fundamental misunderstanding of secure password storage practices.

Uniform Passwords: The use of a single password for all users in a county is another glaring security issue. It creates a single point of failure, making the system vulnerable to unauthorized access if the common password is compromised.

In conclusion, the security practices observed in the **`TabulatorUser`** table indicate a concerning lack of basic cybersecurity protocols, potentially jeopardizing the integrity of the electoral process. The subsequent sections will delve deeper into the implications of these security flaws and the necessary steps to mitigate them.



Assessing the **AppUser** Table

Continuing the exploration of the electoral system's database, I next examined the `AppUser` table. This table, found to be consistent across all four counties, offered a different perspective on the system's approach to password security compared to the previous tables I analyzed.

Table Structure and Findings:

The **AppUser** table listed seven users, each with administrative access. Notably, their passwords were stored as hashes, indicating a correct implementation of password storage protocols, unlike the reversible encryption observed in the **TabulatorUser** table.

username	password (hashed)	firstName	lastName
county	0x82D99B…52EA7B	Bibb	County
Techadvisor	0x6B69EC…ECDFC2	State of	Georgia
MRO01	0x6166A7…986384	MRO	M01
…	…	…	…

The usage of hashed passwords is a positive aspect, demonstrating an understanding of proper security practices for password storage. However, a deeper analysis revealed a critical flaw.

Critical Security Concerns:

Repeated Hashes Indicate Shared Passwords: Despite the proper use of hashing, there were only three different hashes present for the seven users. This implies that multiple administrative accounts shared the same password. In cybersecurity, each user should have a unique password to prevent widespread access in case one password is compromised.

Statewide Uniform Administrative Passwords: A more alarming observation was that for every county, the administrative passwords were identical. This uniformity across the state meant that access to one of these administrative accounts, such as **SAdmin**, **MRO01**, or **ROAdmin**, would grant access to the administrative settings of all the counties. Such a setup poses a significant security risk, as it creates a single point of failure that could potentially lead to widespread system compromise.

The findings in the **AppUser** table, while initially appearing more secure due to the use of hashing, still reveal a lack of fundamental cybersecurity standards. The use of shared passwords, especially for administrative accounts with broad access, is a critical vulnerability. This discrepancy between the security measures applied in different tables and the shared passwords across counties raises questions about the consistency and effectiveness of the system's overall cybersecurity strategy.

The X509 Certificate Data

The discovery of the X509 data certificate in the `ElectionEvent` table of the electoral system's database is perhaps one of the most concerning findings in this investigation, both due to its implications and the complexity involved in understanding these implications.

Understanding Security Certificates:

A security certificate, often issued by a trusted certificate authority, is integral to maintaining the confidentiality and integrity of digital communications. It includes a pair of cryptographic keys: a public key and a private key. The private key must be zealously guarded, as its exposure compromises the security of the entire certificate. If it becomes exposed, immediate revocation and reissuance of the certificate are essential. The public key, in contrast, is freely distributed to facilitate secure interactions.

The Compromise in the Database:

In this particular database, the entire X509 certificate — both the public and private keys — was stored as a plaintext string. This is a critical security flaw. Converting this string into a byte array and assigning it a certificate file extension are all that is required to exploit this vulnerability. Once the certificate file is created, a simple double-click installs it into the user’s certificate store, without any password protection or additional security measures.

Potential Uses of X509 Certificates in an Election System:

X509 certificates play several vital roles in securing digital systems:

Secure Email Communications: They are used to encrypt and digitally sign emails, ensuring that sensitive information remains confidential and verifying the sender’s identity.

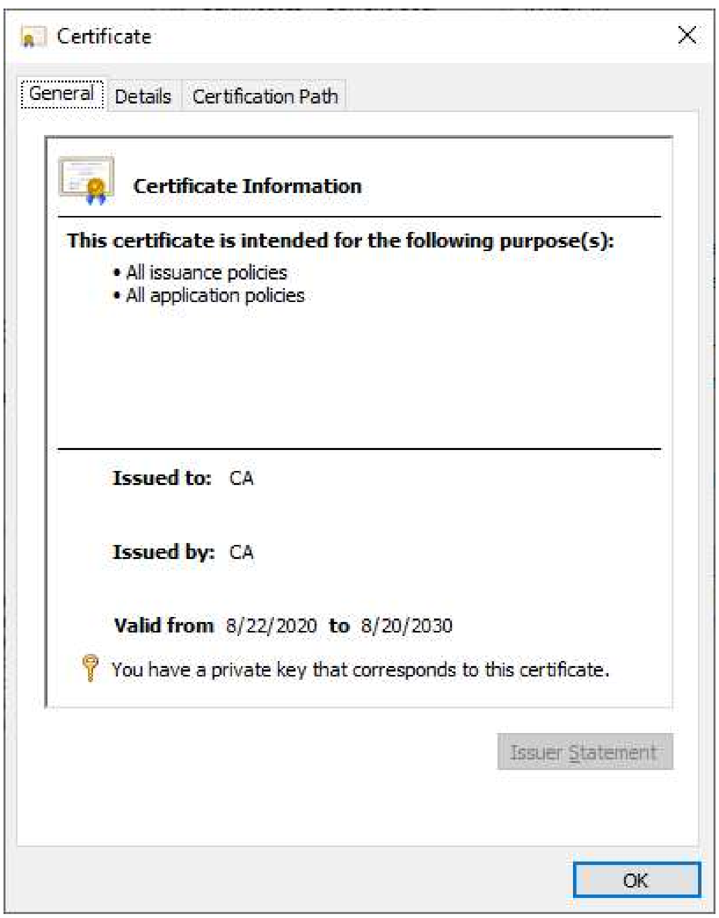
Digital Signatures on Files: Certificates can be used to apply verifiable digital signatures to files. In an election system, this might include voter rolls, election results, or software updates, ensuring the integrity and authenticity of these files.

Encrypted Data Transmission: They are fundamental in setting up SSL/TLS connections, which are crucial for secure data transmission over the internet. This could be used for transmitting election results or voter information between different entities.

Authentication and Access Control: Certificates can play a role in authenticating users or systems before granting access to sensitive data or operations, potentially including access to voting machines or electoral databases.

The Gravity of the Situation:

Leaving the complete X509 certificate, including the private key, unprotected and easily accessible in the database, is akin to leaving the keys in the ignition of an unlocked car. It exposes the electoral system to a multitude of risks, from unauthorized access and data manipulation to the potential fabrication of fraudulent digital communications or transactions. This vulnerability cannot be overstated and calls for immediate and decisive action to secure these systems.



Ballots on Demand

Discovery of 'Fuzzy' Ballots

During the comparison of original and recount ballots in Fulton County, a peculiar pattern emerged. Certain ballots exhibited consistently light or dark un-filled circles, a feature independent of the tabulator or scanner used. This consistency suggested that the characteristic was inherent to the ballots themselves, not a result of scanner variability. These ballots were colloquially termed 'fuzzy' ballots, a phenomenon unique to Fulton County, and appeared in a significant number.

The 'Fuzzy' Ballot Phenomenon: A Theory

My background in scanners and printers led me to theorize about the cause of these 'fuzzy' ballots. Typically, professional print shops use high-quality black ink for printing ballots. However, standard printers, like those available at retail stores, often create black ink by combining colors like Cyan, Magenta, and Yellow. This practice, commonly seen in brands like HP, results in a different shade of black, one that necessitates more frequent ink purchases.

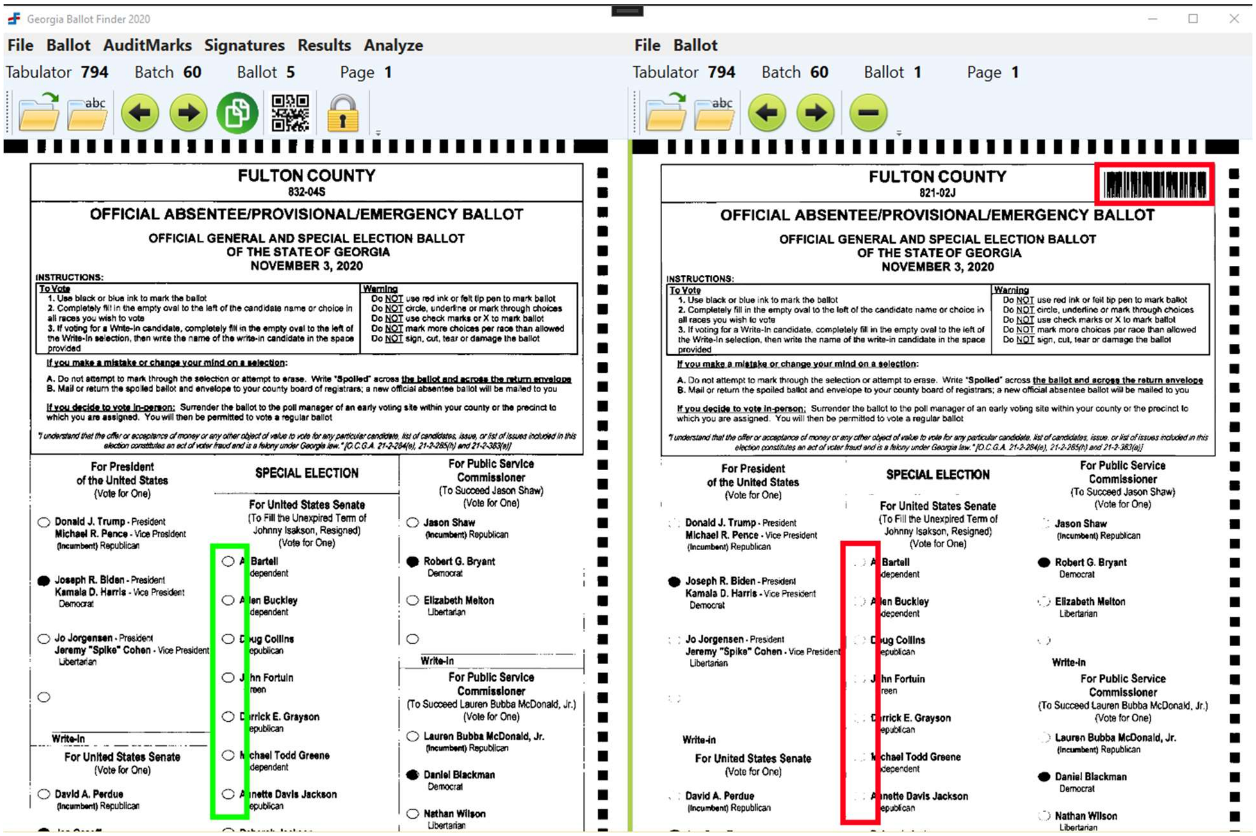
When scanners designed to ignore red ink are used, the non-standard black ink of these locally printed ballots doesn't register as effectively. This results in the faded appearance of the circles, a key characteristic of the 'fuzzy' ballots.

Identification of On Demand Ballots

A critical realization was the presence of a prominent barcode at the top of these ballots. This barcode was indicative of 'On Demand Ballots,' a contingency method used when a county faces a shortage of standard ballots for advance voting.

Inquiries into the prevalence of On Demand Ballots in Fulton County suggested that approximately 5,000 such ballots were generated. With the latest software update, I was able to identify 3,575 of these On Demand Ballots that were cast in the election.

The identification and analysis of On Demand Ballots in Fulton County reveal an important aspect of the voting process, particularly in situations of ballot shortage. The unique characteristics of these ballots, primarily their 'fuzzy' appearance due to printing methods, underscore the nuances in ballot creation and processing.



The Strange Timing of On Demand Ballots

With the On Demand Ballots now identified, a crucial aspect of our analysis focused on the timing of their scanning. This factor provides insight into the distribution and handling of these ballots in the voting process.

Contrary to expectations, these On Demand Ballots did not exhibit a random distribution among early voting ballots. Instead, they tended to appear in batches, often grouped together, which deviates significantly from the expected norm. Under typical circumstances, if these ballots were mailed, received, and processed like other early votes, they would be randomly intermixed with the rest of the ballots.

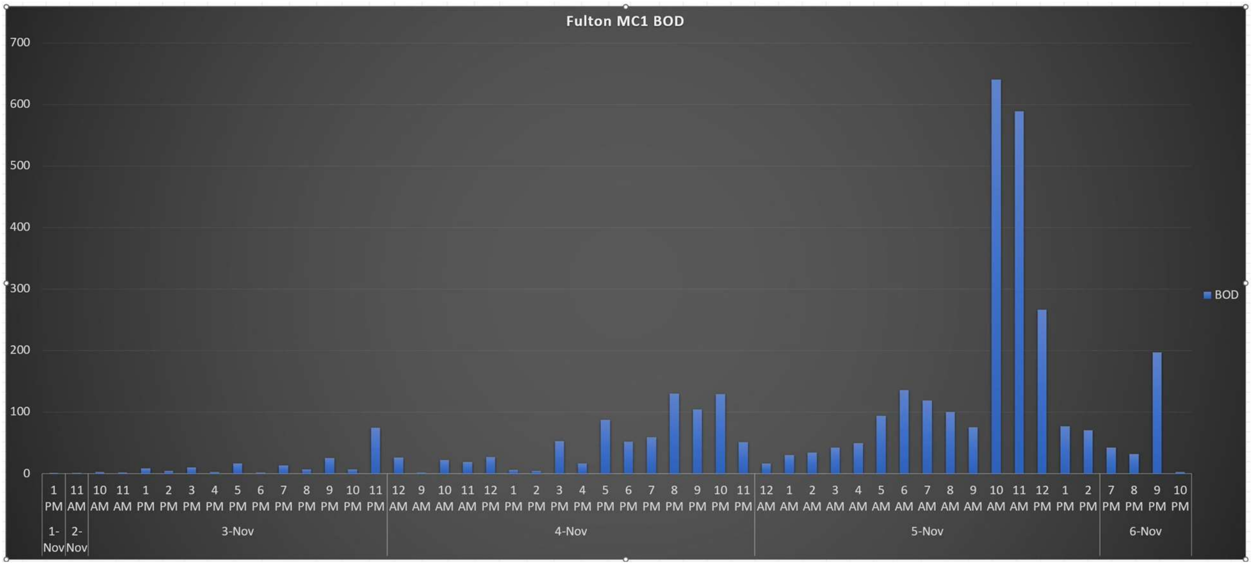
Case in Point: Tabulator 5162 Batch 387

A striking example of this pattern was observed in Tabulator 5162, Batch 387, where 75 consecutive ballots were all On Demand. Such a sequence starkly contradicts the probability of a random mix that would occur through normal postal return processes.

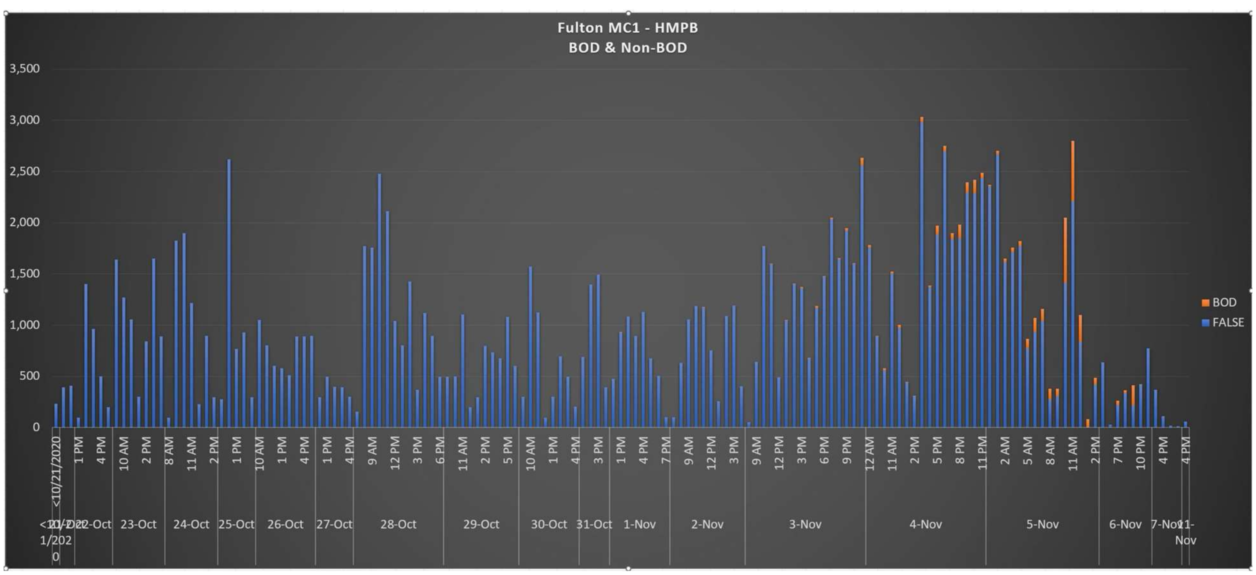
This pattern of grouping was not an isolated incident. In most instances, On Demand Ballots were not randomly dispersed within the batches as one would expect. Instead, they were commonly found clustered together.

Analyzing Scanning Times: An Anomaly

The most intriguing aspect emerged when we analyzed the scanning times of these ballots. On Election Day, only a minimal number of On Demand Ballots were processed. The following day, November 4th, saw a slightly higher but still small number. However, at 10 AM on November 5th, two days post-election, a surge occurred with over 1,800 of these ballots being scanned – an anomaly in the context of time analysis.



Upon charting the scan times of all ballots, it became evident that the On Demand Ballots were predominantly scanned towards the end of the entire ballot scanning period. This finding is highly unusual, given that these ballots, by their nature, should have been randomly distributed among the overall ballot pool.



The peculiar timing and grouping patterns of the On Demand Ballots in our analysis raise significant questions about their handling and processing. These ballots' concentrated appearance in batches, and

their late scanning times, particularly days after the election, point to an irregularity that deviates from standard electoral procedures. This abnormality warrants further investigation to understand the implications and reasons behind such a distribution pattern.

Other Security Issues

The SHA Signature Files and Ballot Image Security

While the integrity of the database is crucial, the security of ballot images holds equal significance. During our review of the data obtained through Open Records Requests, we encountered 'sha' signature files associated with these ballot images. These files are essentially hashes of the ballot images, serving to verify the authenticity and integrity of these images.

Understanding the Ballot Images:

Ballot images in this context are multi-page TIF files. The first two pages represent the front and back of the paper ballot, while the third page is a computer-generated audit mark by the ballot software. After a ballot is scanned and processed, this audit mark is created, followed by the generation of a SHA signature. Theoretically, the usage of hash signing is an adequate method to ensure that the ballot image remains unaltered, as anyone can later regenerate the hash and verify that it matches the original sha file.

The Security Flaw in the Hashing Method:

However, there's a significant flaw in the method employed to create these hashes. The system uses a basic hashing method that does not involve a public/private key pair, which is a hallmark of modern, secure hashing algorithms. This oversight means that anyone could upload a ballot image to one of many online hash generator websites and regenerate the sha signature file. Consequently, there's no reliable method to detect if a malicious actor has altered a ballot and created a new sha file to match.

This ease in recreating sha files creates a false sense of security. It suggests that the ballot images are secure and unaltered, which might not be the case. A ballot can be tampered with, a new sha file generated, and the files would still appear legitimate.

The Adjudication Process and its Impact:

Interestingly, when a ballot undergoes adjudication, the resulting sha signatures no longer match the original ballot. The adjudication process involves adding additional text to the third page of the TIF image, thus altering its hash signature. This discrepancy highlights the system's vulnerability to alterations.

Recommended Best Practices:

To enhance security, a more robust approach would be to employ a hashing algorithm that requires a public/private key certificate. This certificate would generate a unique signature that couldn't be replicated by anyone without access to the private key. The public part of the certificate would then be used to verify the authenticity of the hash. Such an approach would significantly bolster the security of the ballot images, ensuring that any alterations can be reliably detected and traced back to their source.

Mismatch between Ballot Data and the Database

In the complex ecosystem of election technology, the alignment between different components is critical. One area where this alignment is paramount, yet sometimes faltering, is in the relationship between the QR codes on Ballot Marking Device (BMD) ballots and the corresponding database records.

The Technicality of BMD Ballots:

BMD ballots, in their essence, simplify the voter's choice to a binary code. For instance, in the 2020 Presidential election, different binary values represented different candidates: 0x8 for Trump, 0x4 for Biden, 0x2 for Jo Jorgensen, and 0x1 for a write-in vote. This binary system, while logically sound, predicates its accuracy on a perfectly aligned database between the ballot marking devices and the tabulators.

The 2022 DeKalb County Commissionaires Race:

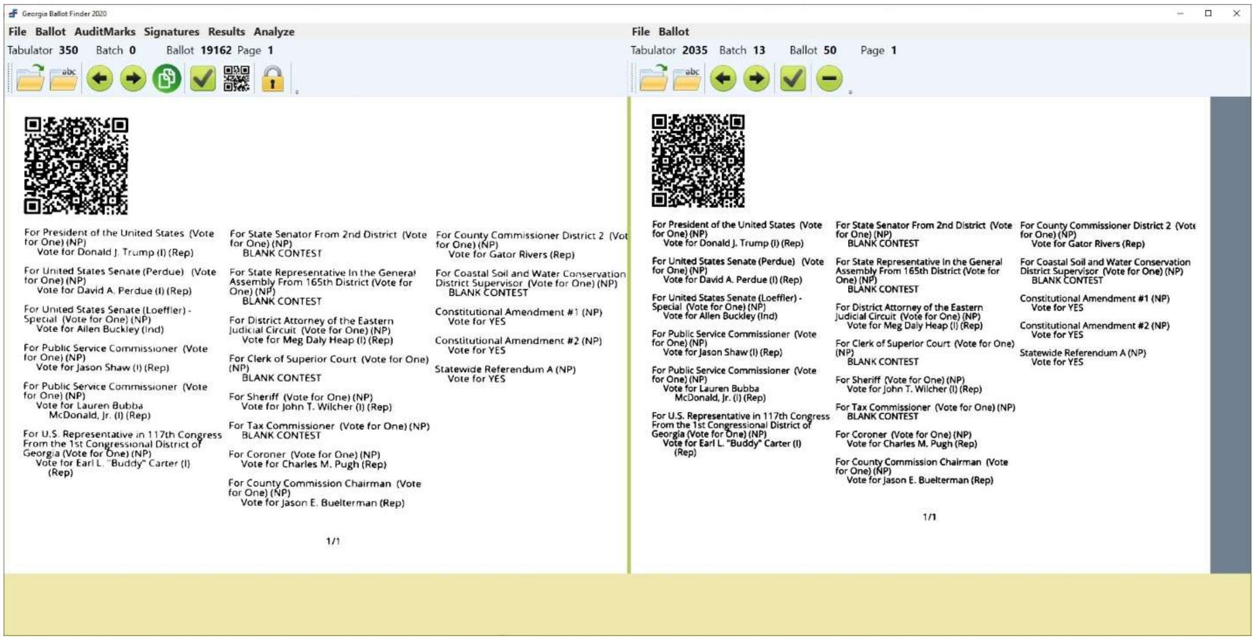
A notable incident in DeKalb County, Georgia, in 2022, illustrates the potential pitfalls of this system. Michelle Long Spears, a candidate in the Commissioner's race, discovered anomalies in her vote counts. In certain precincts, she received no votes, which prompted a deeper investigation. It turned out that votes cast for her were being incorrectly attributed to another candidate, Marshall Orson. This error was traced back to a last-minute candidate withdrawal, which led to a reshuffling of the ballot order. Consequently, a vote for Spears (the fourth bubble) was misinterpreted as a vote for Orson by the tabulator, exposing a critical synchronization issue between the ballot devices and the tabulating system.

Similar Occurrences and Implications:

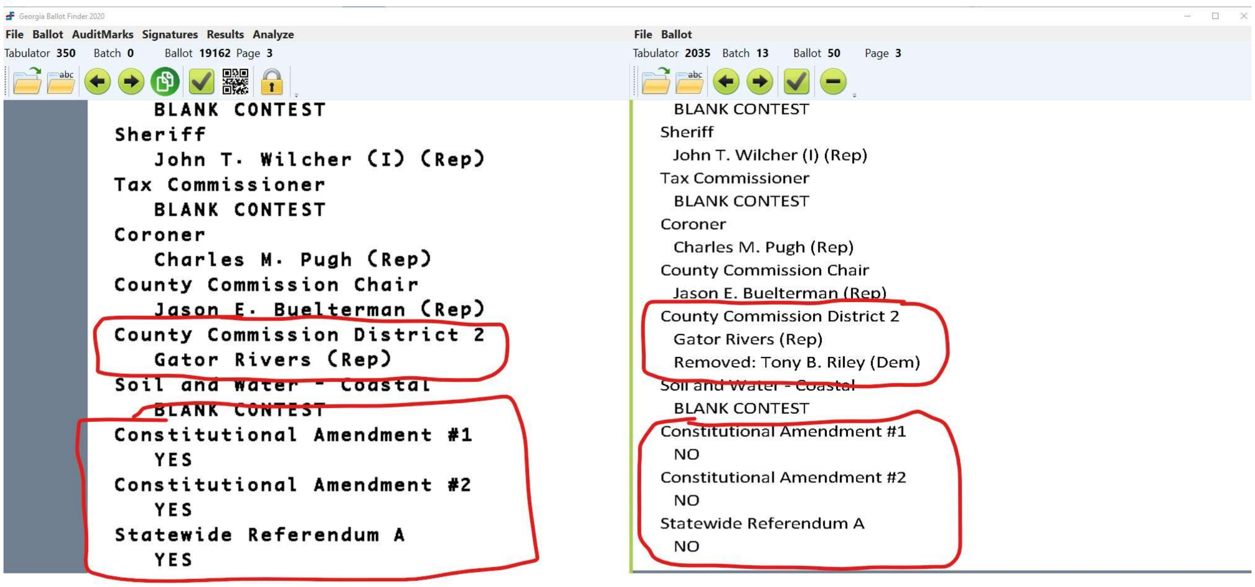
We observed a similar scenario in another county, where the removal of a candidate named 'Tony Riley' between the original vote and the recount affected down-ballot choices. Votes that were intended as 'YES' were altered to 'NO'. Thankfully, since the recount focused solely on the presidential election, this discrepancy did not impact the final outcome. However, it highlights the fragile nature of BMD QR-coded ballots and the imperative need for constant synchronization of the database throughout the counting process.

Conclusion:

This section underscores the vulnerability inherent in the use of BMD ballots and emphasizes the need for database management to ensure accuracy in reflecting voters' intentions. Any misalignment, as shown in these cases, can lead to significant misrepresentations of voter choices, potentially undermining the integrity of the electoral process.



Georgia Ballot Finder 2020
File Ballot AuditMarks Signatures Results Analyze
Tabulator 350 Batch 0 Ballot 19162 Page 1
For President of the United States (Vote for One) (NP) Vote for Donald J. Trump (I) (Rep)
For United States Senate (Perdue) (Vote for One) (NP) Vote for David A. Perdue (I) (Rep)
For United States Senate (Loeffler) - Special (Vote for One) (NP) Vote for Allen Buckley (Ind)
For Public Service Commissioner (Vote for One) (NP) Vote for Jason Shaw (I) (Rep)
For Public Service Commissioner (Vote for One) (NP) Vote for Lauren Bubba McDonald, Jr. (I) (Rep)
For U.S. Representative in 117th Congress From the 1st Congressional District of Georgia (Vote for One) (NP) Vote for Earl L. "Buddy" Carter (I) (Rep)
For State Senator From 2nd District (Vote for One) (NP) BLANK CONTEST
For State Representative In the General Assembly From 165th District (Vote for One) (NP) BLANK CONTEST
For District Attorney of the Eastern Judicial Circuit (Vote for One) (NP) Vote for Meg Daly Heap (I) (Rep)
For Clerk of Superior Court (Vote for One) (NP) BLANK CONTEST
For Sheriff (Vote for One) (NP) Vote for John T. Wilcher (I) (Rep)
For Tax Commissioner (Vote for One) (NP) BLANK CONTEST
For Coroner (Vote for One) (NP) Vote for Charles M. Pugh (Rep)
For County Commission Chairman (Vote for One) (NP) Vote for Jason E. Buelterman (Rep)
For County Commissioner District 2 (Vot for One) (NP) Vote for Gator Rivers (Rep)
For Coastal Soil and Water Conservation District Supervisor (Vote for One) (NP) BLANK CONTEST
Constitutional Amendment #1 (NP) Vote for YES
Constitutional Amendment #2 (NP) Vote for YES
Statewide Referendum A (NP) Vote for YES
1/1
File Ballot
Tabulator 2035 Batch 13 Ballot 50 Page 1
For President of the United States (Vote for One) (NP) Vote for Donald J. Trump (I) (Rep)
For United States Senate (Perdue) (Vote for One) (NP) Vote for David A. Perdue (I) (Rep)
For United States Senate (Loeffler) - Special (Vote for One) (NP) Vote for Allen Buckley (Ind)
For Public Service Commissioner (Vote for One) (NP) Vote for Jason Shaw (I) (Rep)
For Public Service Commissioner (Vote for One) (NP) Vote for Lauren Bubba McDonald, Jr. (I) (Rep)
For U.S. Representative in 117th Congress From the 1st Congressional District of Georgia (Vote for One) (NP) Vote for Earl L. "Buddy" Carter (I) (Rep)
For State Senator From 2nd District (Vote for One) (NP) BLANK CONTEST
For State Representative In the General Assembly From 165th District (Vote for One) (NP) BLANK CONTEST
For District Attorney of the Eastern Judicial Circuit (Vote for One) (NP) Vote for Meg Daly Heap (I) (Rep)
For Clerk of Superior Court (Vote for One) (NP) BLANK CONTEST
For Sheriff (Vote for One) (NP) Vote for John T. Wilcher (I) (Rep)
For Tax Commissioner (Vote for One) (NP) BLANK CONTEST
For Coroner (Vote for One) (NP) Vote for Charles M. Pugh (Rep)
For County Commission Chairman (Vote for One) (NP) Vote for Jason E. Buelterman (Rep)
For County Commissioner District 2 (Vote for One) (NP) Vote for Gator Rivers (Rep)
For Coastal Soil and Water Conservation District Supervisor (Vote for One) (NP) BLANK CONTEST
Constitutional Amendment #1 (NP) Vote for YES
Constitutional Amendment #2 (NP) Vote for YES
Statewide Referendum A (NP) Vote for YES
1/1



Georgia Ballot Finder 2020
File Ballot AuditMarks Signatures Results Analyze
Tabulator 350 Batch 0 Ballot 19162 Page 3
BLANK CONTEST
Sheriff
John T. Wilcher (I) (Rep)
Tax Commissioner
BLANK CONTEST
Coroner
Charles M. Pugh (Rep)
County Commission Chair
Jason E. Buelterman (Rep)
County Commission District 2
Gator Rivers (Rep)
Soil and Water - Coastal
BLANK CONTEST
Constitutional Amendment #1
YES
Constitutional Amendment #2
YES
Statewide Referendum A
YES
File Ballot
Tabulator 2035 Batch 13 Ballot 50 Page 3
BLANK CONTEST
Sheriff
John T. Wilcher (I) (Rep)
Tax Commissioner
BLANK CONTEST
Coroner
Charles M. Pugh (Rep)
County Commission Chair
Jason E. Buelterman (Rep)
County Commission District 2
Gator Rivers (Rep)
Removed: Tony B. Riley (Dem)
Soil and Water - Coastal
BLANK CONTEST
Constitutional Amendment #1
NO
Constitutional Amendment #2
NO
Statewide Referendum A
NO

TITLE V. DISCLOSURES AND DISCOVERY

(iv) Qualifications and Publications: While I have extensive experience in the field of computer science, particularly in fingerprint identification and image analysis, I have not authored any publications concerning ballot analysis in the past 10 years. My expertise in this area is primarily practical, derived from years of hands-on experience and personal research, particularly in developing the 'Ballot Finder' software for ballot analysis.

(v) Expert Testimony in Other Cases: To date, I have not testified as an expert at trial or by deposition in any other cases within the past four years. My involvement in ballot analysis has been independent and not formally linked to any legal cases or proceedings. As far as I am aware, the data and findings from my ballot analysis work have not been utilized in any other legal case."

(vi) Compensation: that I have not received, nor do I expect to receive, any form of compensation for my work. My involvement in developing the 'Ballot Finder' software and conducting ballot analysis has been entirely voluntary, driven by personal interest and commitment to the integrity of the electoral process. Therefore, there will be no financial or other compensation associated with my study and testimony in this case.

I hereby certify that this report is a complete and accurate statement of all my opinions, and the basis and reasons for them, to which I will testify under oath.

Phillip Davis
Jupiter, FL